

VŠB–Technická univerzita Ostrava
Fakulta elektrotechniky a informatiky

DIPLOMOVÁ PRÁCA

2014

Dominik Uhrin

**VŠB–Technická univerzita Ostrava
Fakulta elektrotechniky a informatiky
Katedra telekomunikačnej techniky**

**LIDSKÉ EMOCE A JEJICH ROZPOZNÁVÁNÍ Z REČI
HUMAN EMOTIONS AND THEIR RECOGNITION FROM
SPEECH**

2014

Dominik Uhrin

VŠB - Technická univerzita Ostrava
Fakulta elektrotechniky a informatiky
Katedra telekomunikační techniky

Zadání diplomové práce

Student:

Bc. Dominik Uhrin

Studijní program:

N2647 Informační a komunikační technologie

Studijní obor:

2612T059 Mobilní technologie

Téma:

Lidské emoce a jejich rozpoznávání z řeči
Human Emotions and their Recognition from Speech

Zásady pro vypracování:

1. Emoční stavy a fyziologické projevy emocí.
2. Vytvoření emoční databáze.
3. Návrh a realizace evaluace emocí z řeči subjektivními testy.
4. Vyhodnocení dosažených výsledků a evaluace věrohodnosti databáze.

Seznam doporučené odborné literatury:

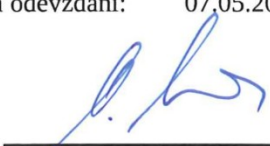
ATASSI, H., SMEKAL, Z., ESPOSITO, A. *Emotion Recognition from Spontaneous Slavic Speech*. In Proceedings of 3rd IEEE International Conference on Cognitive Infocommunications. 2012. pp. 389-394.
PARTILA, P., VOZNAK, M., MIKULEC, M., ZDRALEK, J. *Fundamental Frequency Extraction Method using Central Clipping and its Importance for the Classification of Emotional State*. Advances in Electrical and Electronic Engineering, Volume 10, Issue 4, 2012, pp. 270 – 275.

Formální náležitosti a rozsah diplomové práce stanoví pokyny pro vypracování zveřejněné na webových stránkách fakulty.

Vedoucí diplomové práce: **doc. Ing. Miroslav Vozňák, Ph.D.**


Datum zadání: 01.09.2013

Datum odevzdání: 07.05.2014



doc. Ing. Miroslav Vozňák, Ph.D.
vedoucí katedry



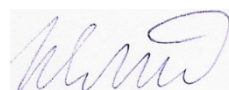


prof. RNDr. Václav Snášel, CSc.
děkan fakulty

Čestné prehlásenie

Prehlasujem, že túto prácu som vypracoval samostatne a uviedol som všetky použité literárne
pramene a publikácie, z ktorých som čerpal.

V Ostrave dne: 06.05.2014



Podpis

PodĎakovanie

Ďakujem svojmu vedúcemu, doc. Ing. Miroslavovi Vozňákovi, Ph.D. za to, že mi pomohol pri
vypracovaní diplomovej práce a poskytol konzultácie a cenné rady. Ďakujem aj za trpezlivosť
svojej rodiny a blízkych, a za to, že mi pomáhali v tomto ťažkom boji. Ďakujem aj katedre za
poskytnutie potrebných prístrojov, pomôcok a materiálov pri realizácii diplomovej práce.

ABSTRAKT

Táto diplomová práca priblíži čitateľovi tematiku emočných stavov a ľudskej reči ako jedného z fyziologických prejavov emočného stavu. Ďalej táto práca chce priblížiť čitateľovi tvorbu databázy emočných vzoriek v podobe krátkych zvukových nahrávok a následne tvorbu nástroja, ktorý pomôže vyhodnotiť vytvorenú databázu pomocou subjektívnej metódy.

Prvá časť sa venuje emočným stavom, konkrétne popisovaniu jednotlivých druhov emócií. Táto časť pojednáva aj o fyziologických prejavoch spojených s prežívaním emócií. Vzhľadom na to, že cieľom práce je vytvoriť databázu zvukových nahrávok, tak sa v tejto časti pojednáva o ľudskej reči, keďže je jednou z fyziologických prejavov. Zaoberá sa nielen tvorbou hlasu, ale aj variantmi reči a jej vlastnosťami. Druhá časť pojednáva o tvorbe už spomínanej databázy zvukových nahrávok, resp. aké požiadavky má databáza spĺňať, aké vlastnosti majú mať zvukové nahrávky. V tretej časti sa venujem tvorbe nástroja na subjektívne hodnotenie databázy emočných vzoriek, ktorú som vytvoril. V poslednej časti vyhodnocujem výsledky dosiahnuté hodnotením vzoriek emočnej databázy a redukujem vzorky databázy na základe analýzy vierohodnosti jednotlivých vzoriek. V závere sa venujem zhodnoteniu celej tejto práce a jej prínosu, resp. možnosti ďalšieho pokračovania v tejto práci.

KLÚČOVÉ SLOVÁ

emócia, emočný stav, reč, hlas, databáza, emoDB, vzorky, emočné vzorky, hodnotenie

ABSTRACT

This diploma thesis is introducing to topic of emotional states and human speech as one of physiological utterance. This topic also introducing reader to creating of database of emotional samples in form of short sound records and mainly creation of tool which is helping in evaluation of created database using subjective method.

First part of thesis is dedicated to emotional states, mainly describing every type of emotion. This part also discusses about physiological utterances linked with emotional experiencing. Considering that the goal of this thesis is creation of sound sample database so this part is dedicated to human speech because it is one of physiological utterances. It is dealing not only with voice creation but also with speech variants and it's features. Second part discusses about creation of already mentioned database of sound samples respectively what goals database has to be met and what features sound records have to have. Third part I am dedicating to creation of tool for subjective evaluation of emotional sample database which I created. In last part of the thesis I am evaluating achieved results of sample evaluation of emotional database and also reducing number of database samples on basis of analyse of individual sample credibility. Conclusion is dedicated to review of the whole thesis a it's asset respectively opportunities of further continuation in this topic.

KEY WORDS

emotion, emotion state, speech, voice, database, emoDB, samples, emotional samples, evaluation

ZOZNAM POUŽITÝCH SYMBOLOV A SKRATIEK

F_0	frekvencie základného tónu reči
T_0	perióda základného tónu reči
Hz	hertz, základná jednotka frekvencie
kHz	tisíc násobok základnej jednotky frekvencie
mm	milimeter, základná jednotka dĺžky
dB	základná jednotka hladiny intenzity zvuku
ms	milisekunda
min	minimálna hodnota
max	maximálna hodnota
F_1 - F_3	frekvencie prvých troch formantov
f_s	vzorkovacia frekvencia, z anglického výrazu <i>sample</i>
n_{formant}	n -tý formant v poradí
MFCC	Melfrekvenčné cepstrálne koeficienty, z anglického výrazu <i>Mel-frequency cepstral coefficients</i>
LTASS	Dlhodobé priemerné spektrum, z anglického výrazu <i>Long time average spectrum</i>
LPCC	Lineárne predpovedný cepstrálny koeficient, z anglického výrazu <i>Linear predictive cepstral coefficient</i>
LSF	Lineárna spektrálna frekvencia, z anglického výrazu <i>Linear spectral frequency</i>
Jitter _{REL}	relatívna hodnota jitteru
Log	logaritmus, inverzná funkcia k exponenciálnej funkcii
T_i	perióda pri jitteri
HNR	pomer harmonickej frekvencie k šumu, z anglického výrazu <i>Harmonic to Noise Ratio</i>
A_i	amplitúda pri shimmeri
ECG	elektro-kardiografia
EEG	elektro-encefalografia
EOG	elektro-okulografia
ECoG	elektro-kortikografia
MEG	magneto-ecefalografia
PET	pozičná emisná tomografia
fMRI	funkčná magnetická rezonancia
fNIR	funkčné blízko infračervenej frekvencie snímkovanie
%	symbol percenta
.wav	koncovka zvukového formátu <i>wave</i> používajúceho PCM kódovanie
.mp3	koncovka pre zvukový formát MPEG vrstvy 3 stratového formátu
.flv	koncovka pre flashové video
PCM	pulzová kódová modulácia
n	skratka pre neutrálny emočný stav
a	skratka pre hnev ako emočný stav z anglického výrazu <i>anger</i>
h	skratka pre šťastie ako emočný stav z anglického výrazu <i>happiness</i>
s	skratka pre smútok ako emočný stav z anglického výrazu <i>sadness</i>
f	skratka pre strach ako emočný stav z anglického výrazu <i>fear</i>
m	skratka pre muža ako označenie pohlavia z anglického výrazu <i>male</i>
f	skratka pre ženu ako označenie pohlavia z anglického výrazu <i>female</i>
emoDB	skratka pre výraz emočná databáza
kB/s	jednotka dátového toku kilobity za sekundu

OBSAH

ÚVOD.....	8
1 EMOČNÉ STAVY A ICH FYZIOLOGICKÝ PREJAV	9
1.1 Emócia	9
1.1.1 Ako vzniká emócia.....	10
1.1.2 Význam emócie	11
1.1.3 Emočné stavy	12
1.1.4 Radosť	12
1.1.5 Neutrálny stav	13
1.1.6 Nuda	13
1.1.7 Smútok	14
1.1.8 Hnev	14
1.1.9 Strach	14
1.1.10 Znechutenie	15
1.2 Ľudská reč.....	15
1.2.1 Hlasový trakt	15
1.2.2 Základný tón reči.....	16
1.2.3 Vznik formantov	17
1.2.4 Varianty reči	17
1.3 Vlastnosti reči.....	18
1.3.1 Prozodické vlastnosti	18
1.3.2 Spektrálne vlastnosti.....	19
1.3.3 Vlastnosti hlasovej kvality	20
1.3.4 Prozodické vlastnosti českého jazyka.....	20
1.4 Objektívne metódy pre určovanie emócií.....	21
2 VYTVÁRANIE EMOČNEJ DATABÁZY	22
2.1 Požiadavky na databázu	22
2.2 Pomenovanie nahrávok emočnej databázy	22
2.3 Parametre nahrávok emočnej databázy	23
2.4 Zdrojové nahrávky.....	25
2.5 Spracovanie zdrojových záznamov	26
3 NÁVRH A REALIZÁCIA HODNOTENIA EMÓCIÍ.....	28
3.1 Voľba subjektívnej metódy hodnotenia.....	28

3.2 Príprava pre návrh a realizáciu nástroja	29
3.3 Návrh vzhľadu a štruktúra nástroja.....	30
3.4 Návrh databázy emoDB.....	31
3.5 Návrh algoritmu nástroja	33
3.5.1 Algoritmus uvítacej stránky nástroja	33
3.5.2 Algoritmus hodnotiacej stránky nástroja.....	35
3.5.3 Algoritmus ostatných stránok nástroja	39
4 VYHODNOTENIE VÝSLEDKOV A VIEROHODNOSTI DATABÁZY	40
4.1 Vyhodnotenie dosiahnutých výsledkov databázy.....	40
4.2 Určenie vierohodnosti databázy emočných vzoriek	42
ZHODNOTENIE A ZÁVER.....	46
LITERATÚRA A ODKAZY	47
ZOZNAM PRÍLOH	48

ÚVOD

ÚVOD

Vývoj aplikácií a služieb sa snaží o prínos prirodzenej interakcie medzi človekom a digitálnou prístrojmi, ako sú smartfóny, televízie osobné počítače a podobne. Špecifikovanie príkazov pomocou hlasu alebo gesta sú veľmi populárne v dnešnej dobe. Značné množstvo informácií je získaných z ľudskej reči s celkom veľkou presnosťou. Ľudská reč tiež obsahuje sekundárne informácie, ktoré charakterizujú rečníka a jeho stav. Sú to napríklad vek, pohlavie, emočný stav alebo odchýlky v reči, ako napríklad šušľanie, resp. ďalšie vlastnosti, ktoré sú obsiahnuté v ľudskej reči. Tieto parametre môžu byť ovplyvnené rečníkom, ktorý má ťažkosti. Preto tieto sekundárne informácie ľudskej reči sú zviazané s dôveryhodnými informáciami.

Ako naznačuje názov, táto práca opisuje emočné stavy a nástroj na získanie emočného stavu pomocou subjektívnej metódy. Emócia predstavuje jednu z charakteristických vlastností človeka, ktoré opisujú jeho mentálny stav, ktorý je ovplyvňovaný fyziologickými zmenami prebiehajúcimi v ľudskom tele. Tieto zmeny sa tiež odrážajú na ľudskej reči. Informácie o emočnom stave sú vyžadované vo viacerých oblastiach. Štatistické ohodnotenie spokojnosti zákazníka a jeho záujem o produkty sú hodnotené na základe emočného stavu. Tieto informácie sú priamou odpoveďou na nejaký stimul. Agenti v call centrách môžu byť ohodnotení na základe ich vykonanej práce a ich prístupu ku zákazníkom. Je tu možnosť vytrénovať nových agentov a naučiť ich správnym komunikačným zručnostiam.

Ľudské telo ovplyvnené silnými emóciami sa dostáva do stresu. Pozitívny stres je zapríčinený emóciou ako šťastie, napríklad pri výhre v lotérii. Negatívny stres je naopak zapríčinený psychologickým tlakom, napríklad strach z toho, že nestihnem odovzdať túto prácu.

Štátne sektory ako polícia, požiarnici a hlavne armáda vytvára najväčší emočný nátlak na svojich zamestnancov. Schopnosti vojaka, policajta alebo požiarnika vykonávať prácu, záleží na pokynoch, ktoré dostávajú z dispečingu. Použitie systému na rozpoznávanie emočného stavu človeka z reči znemožňuje zmeniť príkazy z dispečingu pre muža v akcií. Systém pre rozpoznávanie emócií prináša možnosť riadiť príkazy a ďalšie kroky pre muža v akcií.

Aplikácie vyžadujúce bezpečnosť často používajú reč ako autorizačný kľúč. Rečový prejav je ovplyvňovaný fyziologickými zmenami, ktoré sú zapríčinené zmenami emočného stavu. Aj autorizovanej osobe môže byť zamietnutý prístup, pretože autorizačná časť rozpoznala zmeny v rečovom prejave a vyhodnotila ho ako zlý kľúč. Preto je užitočné používať sekundárne informácie z reči, ktoré môžu zabrániť falošným zákazom prístupu. To je len niekoľko možností ako sa môže používať rozpoznávanie emočného stavu človeka.

Prejdem k zámeru mojej práce. Tá bola vypracovaná za účelom vytvorenie skupiny vzoriek, v podobe zvukových nahrávok, ktorá mala byť následne ohodnotená skupinou subjektov. Na základe týchto hodnotení sa mal zistiť emočný stav ľudí, ktorých reč bola na nahrávkach zaznamenaná. Na subjektívne hodnotenie bol použitý nástroj, ktorý som vytvoril. Jednalo sa o jednoduchú webovú stránku prepojenú s databázou. Následne som dosiahnuté hodnotenia analyzoval a na základe vierohodnosti vzoriek som vytvoril 3 databázy vzoriek s rôznymi úrovňami vierohodnosti.

1 EMOČNÉ STAVY A ICH FYZIOLOGICKÝ PREJAV

1 EMOČNÉ STAVY A ICH FYZIOLOGICKÝ PREJAV

Vzhľadom k tomu, že prvý bodom zadania je popísať emočné stavy, tak sa najprv v tejto kapitole budem venovať emóciám. Najprv krátko popíšem, čo vlastne tá emócia je a čo predstavuje. Ďalej popíšem význam emócie a teóriu Roberta Plutchika, ktorú v celku dobre popisuje emočná kružnica. V krátkosti popíšem aj jednotlivé základné emočné stavy, ktoré človek dokáže prežívať.

V ďalšej časti sa budem venovať fyziologickému prejavu emočných stavov. Zvolil som si ľudskú reč ako fyziologický prejav emócie, pretože ako druhý bod zadania mám vytvoriť databázu emočných vzoriek vo forme zvukového záznamu. Teda budem mať vytvárať záznamy ľudskej reči v rôznych emočných stavoch. Najprv teda začnem krátkym popisom časti tela, ktoré sa podieľajú na vytváraní hlasu, resp. na vytváraní ľudskej reči. Jedná sa o takzvaný hlasový trakt, ktorý vytvára hlas. Následne prejdem na hlasivky. Táto časť vytvára základný tón hlasu. V ďalšej časti prejdem na formanty. Jedná sa o vedľajšie tóny, ktoré dotvárajú konečný hlasový prejav a bez ktorých by nevznikala ľudská reč.

V predposledných časti spomeniem vlastnosti reči. Vlastnosti môže rozdeliť do troch skupín a to : prozodické, spektrálne a vlastnosti hlasovej kvality. Súčasťou tejto podkapitoly sú aj prozodické vlastnosti českého jazyka. Ako poslednú podkapitolu som uviedol možnosti objektívne metódy zisťovania emočného stavu pozorované subjektu.

1.1 Emócia

Emócia je komplexný jav, ktorý zo sebou nesie zážitok a prejavy v správaní. Výraz emócia všeobecne znamená zmenu psychologického stavu, ktorý sa spúšťa pri dosahovaní určitého cieľa. Ak je cieľ dosiahnutý, môžeme hovoriť o kladnej emócií. Pokiaľ naopak cieľ nie je dosiahnutý môžeme hovoriť o zápornej emócií. [4]

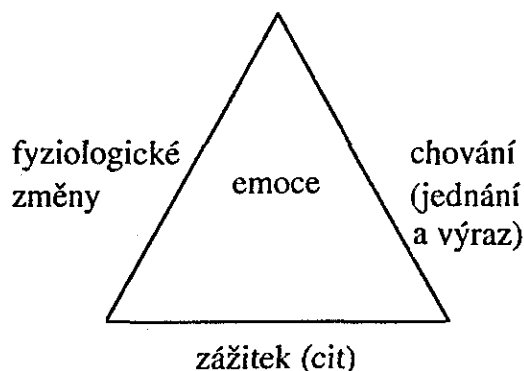
Výraz emócia je spätý s výrazom cit, pretože slovo cit je definované všeobecne ako zážitok; napríklad radosť, smútok, hnev a podobne. Avšak pojem emócie sa môže používať v dvoch významoch:

1. vzťahuje sa k prežívaniu, ako napríklad ak niekto pociťuje úzkosť, ale je vhodnejšie ho označovať výrazom cit;
2. vzťahuje sa nielen k citu, ale aj fyzickému, telesnému stavu a prejavu. [4, 5]

Avšak pojem emócie je významovo nadriadený pojmu citu, ktorý môže byť chápaný ako zážitková zložka alebo súčasť emócie. Tieto tri komponenty, ktoré vytvárajú samotnú emóciu môžeme vyjadriť nasledujúcou schémou na obrázku 1.1. [5]

Emócia sa prejavuje hlavne u ľudí, ktorý nemajú problém s prejavom emócie. Resp. sa prejavuje napríklad u ľudí, ktorí sa nachádzajú v nečakanej alebo neznámej situácii, teda na tieto situácie nemajú nejaký zabehnutý vzorec, podľa ktorého by sa chovali. Aj preto deti prejavujú viacej druhov emócií ako dospelí ľudia. Samozrejme rokmi si deti osvoja určité vzorce správania, podľa ktorých sa budú prejavovať. [5]

1 EMOČNÉ STAVY A ICH FYZIOLOGICKÝ PREJAV



Obrázok 1.1 Schéma zobrazujúca prepojenie jednotlivých komponent tvoriacich emóciu. [5]

Emócia sa neprejavuje len v chovaní, ale aj v reči. Súčasťou rečového prejavu sú aj prozodické zmeny. Dochádza k nim pri prežívaní určitej emócie, sú jej súčasťou. Reč obsahuje približne 10 percent informácií o prežívanej emócií, resp. o stave hovoriaceho človeka. Aj napriek malému množstvu je možné informácie považovať za dôveryhodné, pretože emócia je vývojovo staršia ako rozumové chovanie alebo rozprávanie. Emócia je ľahko prenositeľná z človeka na človeka, ako napríklad plač na pohrebe. [4, 5]

1.1.1 Ako vzniká emócia

Emócie vznikajú spontánne, nie je možné ich vyvolávať umelo. Je možné si však predstaviť určitú situáciu alebo vyvolať si spomienku a tak si privodiť príslušný emočný stav. Emócie v skutočnosti predstavujú ciele chovania, ľudia majú chuť skúšať rôzne príjemnosti, vzrušenia alebo pokoj. Chcú sa zabávať, radovať, ale aj vyhýbať sa pocitu bolesti a utrpenia, resp. smútku, neistoty, nudy a podobne. Ľudia si privolávajú emočný stav takým spôsobom, že sa chovajú tak, aby dosiahli určitého účinku tohto chovania. Ako príklad by som mohol uviesť, že človek si niečo kúpi, aby si urobil radosť alebo si pozrie v televízii nejaký napínavý film. Keď si tak zoberiem, emócie majú vždy reaktívny charakter. Môžu byť vyvolané vonkajšími podnetmi, vzniknutými vnútornými stavmi alebo reakciami, ktoré ukončujú motivované chovanie. Sú tým myslené situácie ako napríklad po jedení človek dosiahne pocit zasýtenia, resp. dovŕšenie pohlavného aktu, ktorý je ukončený orgazmom. Reaktívna povaha emócií je ich hlavným znakom, sú takmer vždy reakciou na vonkajšie podnety a vnútorné stavy. I keď sú sprievodným javom chovania alebo nejakej mentálnej činnosti, napríklad premýšľania, sú opäť reakciou na vlastnosti takého priebehu akým je napríklad jeho plynulosť, vzniknuté prekážky, blížiaci sa úspech a podobne. Emócie vznikajú hlavne vtedy, keď je dané niečo vzrušujúce alebo významné. Vzrušenie a významnosť sú vedľajšie príznaky vzniku emócií. To čo je významné, je súčasne aj vzrušujúce, príkladom uvediem dosiahnutý úspech alebo stretnutie so známou osobnosťou a podobne. [5]

Doposiaľ uvedené ukazuje na to, že emócie nie je možné úmyselne priamo vyvolať ako napríklad pozornosť alebo predstavy. Avšak táto skutočnosť vyvoláva otázku, či je možná úmyselná kontrola emócií, napríklad úmyselné sa zbavovanie strachu, hnevu alebo smútku. Úmyselné zbavovanie sa radosti, to by bolo asi absurdné, ale existujú aj takí ľudia. Neprijemnej emócie sa môžeme zbaviť len, ak ju nahradíme inou emóciou, príjemnou. Samozrejme to nejde

1 EMOČNÉ STAVY A ICH FYZIOLOGICKÝ PREJAV

v každom prípade, ako napríklad zbavovanie sa nepríjemného vnútorného napätia alebo úzkosti autosugesciami. Tento úkon však môže byť náročný a dlhý proces. Častejší je vtieravý charakter niektorých emócií, napríklad hlbokého smútku, ktoré sa nedá zbaviť, pretože človeka pohlcuje a nedovoľuje mu sa sústrediť na niečo iné. Rovnako vtieravé môžu byť city hnevu a nenávisti, ktoré rovnako ako smútok a ďalšie city, môžu človeka zastihnúť v tých najneočakávanejších chvíľach, pretože nie je schopný na sto percent kontrolovať priebeh svojich spomienok, resp. asociácií predstav. Do značnej miery je však možné kontrolovať výraz emócie, to je potlačovanie motorických prejavov emócie a s nimi spojených gestikulácií a mimiky (obe sú zvláštnym prípadom motorických prejavov). Avšak nemôžeme potlačiť sprievodné vegetatívne prejavy, ako je napríklad potenie. Lenže potlačovanie emocionálneho vyjadrovania nebyva dokonalé, zjavujú sa určité pozostatky, napríklad zatínanie zubov a prižmurovanie pri prežívaní hnevu. [5]

Všeobecne je možné rozlišovať vonkajšie a vnútorné podmienky vzniku emócií. Emócie sú vyvolávané významnými vonkajšími situáciami, ako je napríklad prekážka, ktorá stojí v ceste v cielenom jednaní (frustrácia) alebo nebezpečie. Ďalej sú vyvolávané zmenou vnútorných psycho-fyziologických procesov alebo stavov, ako je napríklad hlad, vzniknuté potiaže pri riešení problémov, spomienok a podobne. [5]

L. Schmidt-Atzert rozlišuje zvláštny druh situačných podmienok vzniku emócie, ktorý označuje udalosti bežného života asi preto, že psychológovia venovali prevažne pozornosť veľkým udalostiam, hlavne frustráciám a konfliktom ako typickým emociogenným situáciám. Z týchto veľkých emociogenných situácií alebo udalostí je samozrejme najzávažnejší stres. Zaťažuje situácie, ktorým sa nie je možné prispôbiť, ale ktorej je subjekt z nejakých príčin vystavený, napríklad úloha vyžadujúca veľké psychické alebo telesné vypätie, konfliktný sociálny vzťah a podobne. [5]

1.1.2 Význam emócie

Ako som už vyššie spomenul, emócie sú evolučne staršie ako hovorený prejav. Napríklad malé dieťa, ktoré ešte nevie hovoriť, tak prejavuje svoje potreby napríklad pomocou plaču, keď ho niečo bolí alebo niečo mu chýba. Neskôr samozrejme, keď dieťa vyrastie a naučí sa rozprávať, tak prejavuje svoje potreby aj rečovým prejavom.

Je potrebné si aj uvedomiť, že emócie ľudí sa v priebehu života postupne vyvíjajú, učia sa ako reagovať v nových, neznámych situáciách. Existuje mnoho teórií o tom ako emócie vznikli alebo k čomu ich potrebujeme. [6]

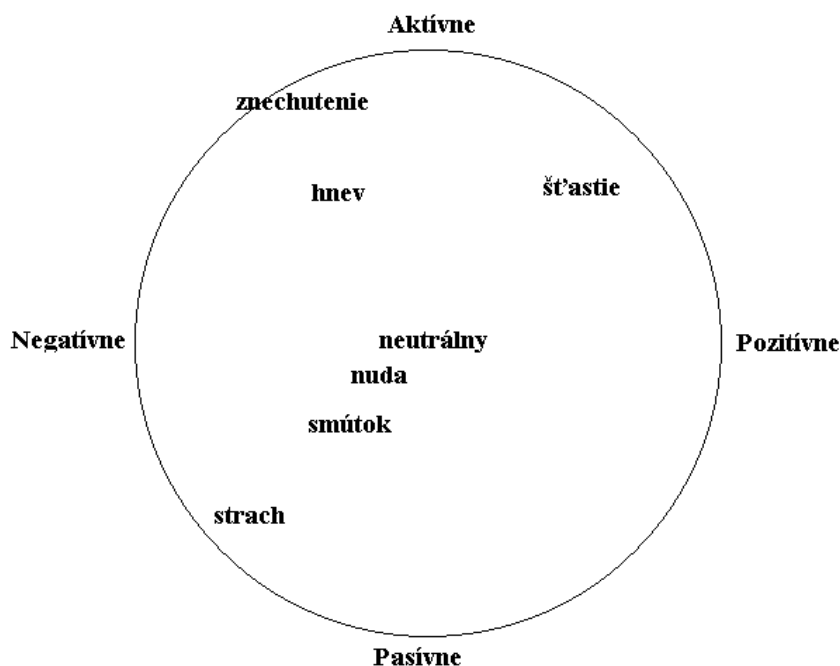
Jednou z mnoho teórií je teória od Roberta Plutchika, ktorá hovorí o tom, že emócie sú vrodené funkcie, ktoré pomáhajú človeku prežiť. Táto teória využíva tzv. emočnú kružnicu, ktorá je rozdelená na negatívnu a pozitívnu časť. V závislosti od toho ako veľmi je emócia prejavovaná a ako veľmi je negatívna alebo pozitívna, podľa toho sa určuje pozícia emócie na v kružnici (pozri obrázok 1.2) . [6]

Iná Cannon-Brandová teória je založená na tom, že pokiaľ človek nachádza v udalosti, ktorá nejakým spôsobom ovplyvňuje jeho emočný stav, tak mozog obdrží správu o tejto udalosti. Tá je následne spracovaná mozgovou kôrou, ktorá je prepojená so samotnými emóciami. Časť zo správy je spracovaná aj hypotalamom, ktorý má na starosti telesné zmeny ako napríklad kontrakcie svalov, tepovú frekvenciu a podobne. [6]

Emócie môžu mať tieto základné úlohy:

1 EMOČNÉ STAVY A ICH FYZIOLOGICKÝ PREJAV

1. reakcia na podnety v prípade, že je nutná okamžitá reakcia
2. ukazovateľ toho, čo máme v pláne v blízkom okamihu urobiť
3. súčasť komunikácie, keď podľa prezentácie myšlienky počas rozhovoru poznáme náladu komunikujúceho
4. vyrovnanie sa s momentom, ktorý práve nastal.



Obrázok 1.2 Emočná kružnica.

1.1.3 Emočné stavy

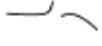

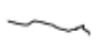
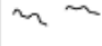




V predchádzajúcej podkapitole som písal o Plutchikovej teórii, ktorá je založená na emočnej kružnici, pozri obr. 1.2. Táto kružnica popisuje 7 emócií, teda 7 stavov, v ktorých sa človek môže nachádzať. Jedná sa o základné stavy, ktoré budem používať aj ja pri vytváraní databázy. Fyzikálne parametre týchto stavov sú znázornené na obrázku 1.3 a konkrétna sa jedná o nasledujúce emočné stavy:

- radosť
- neutrálny
- nuda
- smútok
- hnev
- strach
- znechutenie

1.1.4 Radosť

Všeobecne je radosť veľmi príjemný pocit, spojený s celkovým oživením. Jedná sa o reakciu na úspech alebo zisk. Za slabšiu formu radosti môžeme považovať spokojnosť.

1 EMOČNÉ STAVY A ICH FYZIOLOGICKÝ PREJAV

	neutralita	vztek	smutek	radost	strach	překvapení	obdiv	ironie
průměr F0	120 Hz	180 Hz	110 Hz	200 Hz	120 Hz	120 Hz	150 Hz	190 Hz
min. a max. hodnoty F0 a jejich rozdíl	100 ; 170 E = 70 Hz	100 ; 130 E = 30 Hz	100 ; 130 E = 30 Hz	100 ; 250 E = 150 Hz	100 ; 130 E = 30 Hz	100 ; 340 E = 240 Hz	120 ; 200 E = 80 Hz	100 ; 400 E = 300 Hz
tvar intonační křivky								
průměrný rozdíl intenzity od neutr promluvy	0 dB	11 dB	0 dB	16 dB	0 dB	-1 dB	3 dB	3 dB
min., max. intenzita a rozdíl	+4 ; -2 E = 6 dB	+16 ; +3 E = 13 dB	+2 ; -3 E = 5 dB	+18 ; +8 E = 10 dB	+6 ; -3 E = 9 dB	0 ; -3 E = 3 dB	+10 ; +2 E = 8 dB	+5 ; -3 E = 8 dB
pauza	0 ms	0 ms	0 ms	750 ms	0 ms	250 ms	0 ms	0 ms
trvání věty	1700 ms	1650 ms	1950 ms	2700 ms	1850 ms	2300 ms	1920 ms	1800 ms

Obrázok 1.3 Základné parametre jednotlivých emočných stavov. [6]

Naopak najsilnejšou formou radosti je šťastie. Avšak za vrcholný pocit radosti sa môže považovať extáza. Jedná sa o stav mysle, ktorý odpovedá psycho-fyziologickej reakcií na vyvrcholenie sexuálnej stimulácie. Zdrojom extázy nemusí byť len požitie drogy alebo náboženské osvietenie, môže to byť kľudne postupné dosiahnutie vzrušenia z hudby alebo z krásy prírody. Radosť prežíva človek práve vtedy, keď niečo vytvorí alebo niečo dosiahne. Radosť môže mať človek, pokiaľ hrá hry, zabáva sa alebo vtipkuje. [5]

Výraz radosti je jednoducho rozpoznateľný. Človek ma na tvári charakteristickú krivku úst, lesknúce sa oči, smiech. Fyzické prejavy radosti sú najmä rozšírenie krvných vlások, zvýšenie srdcového tepu a zvýšené prekrvenie (sčervenanie pokožky). Po motorickej stránke sa prejavuje väčším počtom pohybov. [5]

Pokiaľ hovoríme o radosti ako o zvukovom prejave, tak pre tento stav je typická vysoká frekvencia. V priemere sa pohybuje okolo hodnoty 200 Hz. Táto emócia má svoje charakteristické tempo reči. Rečový prejav v tomto stave sa vyznačuje až o 16 dB vyššími hodnotami ako v neutrálnom stave. Radosť patrí medzi aktívne pozitívne emócie, pozri obrázok 1.2. [6]

1.1.5 Neutrálny stav

Keď hovoríme, že človek je v neutrálnom stave, myslíme tým, že momentálne neprežíva žiadnu z emócií. Tento stav môžeme považovať za referenčný stav, keďže sa nevyznačuje žiadnou z emócií, teda ho budeme používať pri porovnávaní s ostatnými emočnými stavmi. Rečový prejav človeka v tomto stave teda neobsahuje žiadne informácie o emócií. Tento stav je umiestnený presne v strede emočnej kružnice a dosahuje priemernej frekvenčnej hodnoty okolo 120 Hz. [6]

1.1.6 Nuda

Jedná sa o stav, ktorý je dosť podobný neutrálnemu stavu. Avšak na rozdiel od neutrálneho stavu je možné zachytiť pri rečovom prejave nejaké informácie o emócií. Tempo rečového

1 EMOČNÉ STAVY A ICH FYZIOLOGICKÝ PREJAV

prejavu býva zvyčajne nižšie. Tento stav má podobný frekvenčný rozsah ako neutrálny stav, avšak jedná sa o mierne pasívnu a negatívnu emóciu, pozri obrázok 1.2. [6]

1.1.7 Smútok

Je to stav rezignácie, v ktorom sa prelína zúfalstvo s momentmi vzdoru a ďalšími pocitmi, ktoré súvisia s týmto stavom. Slabšou formou tohto stavu je zármutok, ktorý môže byť vyvolaný napríklad stratou nejakého predmetu. Najintenzívnejšou formou stavu je žiaľ ako napríklad, keď zomrie osoba, ktorú ste mali radi. Všeobecne môžeme hovoriť, že smútok je vlastne reakciou na stratu niekoho alebo niečoho. Intenzita prežívania tohto emočného stavu závisí na tom, akú hodnotu mala daná stratená vec alebo bytosť. [5]

Smútok je sprevádzaný fyziologickými prejavmi ako poruchy krvného obehu, arytmia srdca, zmeny v rytme dýchania, nechutenstvo, vyčerpanie a poruchy vo vzťahoch s inými osobami. Častým javom v tomto stavu je plač sprevádzaný produkciou slz. [5]

Čo sa týka frekvenčných hodnôt rečového prejavu, je podobný s neutrálnym stavom, avšak rozsah týchto hodnôt je malý, len 30 Hz. Smútok nevykazuje žiadny rozdiel intenzity rečového prejavu vzhľadom na neutrálny stav, teda je na rovnakej úrovni intenzity. Túto emóciu radíme medzi mierne pasívne, ale hlavne negatívne emócie, pozri obrázok 1.2. [6]

1.1.8 Hnev

Hnev všeobecne môžeme chápať ako vrodenu reakciu na prekážku, ktorá bráni dosiahnutiu nejakého cieľa a bráni v konaní. Súčasťou hnevu je aj útok, ktorý má slúžiť na odstránenie nejakej prekážky, aby bolo umožnené ďalšie konanie. Za slabšiu formu tohto stavu môžeme považovať pocit rozhnevania, keď sa človek nahnevá kvôli tomu, že sa mu niečo nedarí uskutočniť. Napríklad keď dopíše pero a ďalšie, ktoré by písalo, nie je k dispozícii. Najsilnejšou formou tohto stavu je amok, resp. záchvat hnevu. Spolu s hnevom je spojená agresia, ktorá môže byť úzko spätá s útokom, za cieľom zničiť prekážku, ktorá bráni v dosiahnutí cieľa. [5]

Stav hnevu sa telesne prejavuje na zvýšení tepu srdca, čo sa prejavuje na rozšírení ciev (hlavne v tvári a na čele) a začervenení, prípadne opuchnutí. Zvyšujú sa kontrakcie svalstva, čo môže zapríčiniť prudké pohyby. Výraz tváre je zamračený a je možné pozorovať zaťaté päsť a zuby a ďalšie mimické prvky. [5]

Jedná sa o aktívny, negatívny stav. Jeho frekvenčný rozsah, počas rečového prejavu, patrí k jedným z najnižších z emočných stavov. Tento stav je charakteristicky relatívne vysokou priemernou frekvenciou základného tónu. Pri intenzite prejavu je rozdiel oproti neutrálnemu stavu značný, až 11 dB, avšak nedosahuje takých hodnôt ako radosť. [6]

1.1.9 Strach

Z všeobecného hľadiska môžeme o strachu hovoriť ako o reakcii na nebezpečie, ktoré práve prebieha alebo, ktoré nastane. Je reakciou na hrozbu bolesti alebo straty života. Podobne ako pri hneve sa jedná o vrodenu reakciu. Najmiernejšou formou strachu je obava, že niečo sa stane. Naopak najsilnejším prejavom strachu je afekt označovaný ako hrôza, zdesenie. Špeciálnym druhom strachu je naučený strach, plynúci z už zažitých nepríjemných momentov, ako napríklad nepríjemná bolestivá návšteva u zubára. Za fyziologické príznaky strachu môžeme považovať znížené vylučovanie slín, vystúpenie studeného potu, zvýšený tep srdca, tzv. husia koža alebo jej zblednutie kože, svalové kontrakcie a ďalšie s tým spojené príznaky. [5]

1 EMOČNÉ STAVY A ICH FYZIOLOGICKÝ PREJAV

Jedná sa o stav, ktorý má podobne frekvenčné hodnoty rečového prejavu ako neutrálny stav. Rozsah frekvenčných hodnôt patrí medzi najmenšie spomedzi stavov. Jedná sa o pasívny emočný stav. [6]

1.1.10 Znechutenie

Všeobecne môžeme o tomto stave hovoriť ako o reakcii na nepríjemnú vec alebo situáciu, ktorá sa práve stala alebo ktorá sa len stane. Je to emočný stav, ktorý slúži na ochranu organizmu pred požitím potenciálne nebezpečných látok a tým sa vyhnúť ochoreniu. Tento emočný stav je typický spájaný s vecami, ktoré vyzerajú špinavé, nepoživatelné, nákazlivé, nechutné alebo nejako inak nepríjemné. Túto emóciu sprevádza zníženie srdcového tlaku a tým aj tepu, zmeny v dýchacom systéme a zníženie vodivosti kože. Pre tento emočný stav je typickým telesným prejavom hlavne mimika tváre. Osoba, ktorá je momentálne v stave znechutenia, má mierne zúžené obočie, mierne zahnuté kútiky úst smerom nadol, zvrásnený nos a na jazyku môže mať viditeľné výčnelky. Opäť podobne, ako u dvoch predchádzajúcich emočných stavov, aj teraz sa jedná o vrodennú emóciu. Avšak na druhej strane je možné tento stav považovať aj za naučený. Napríklad dieťa, keď je ešte malé a po prvýkrát skúša nejaké jedlo, tak mu nemusí chutiť a počas zvyšku života ho už nebude chcieť konzumovať. [7]

1.2 Ľudská reč

Ľudská reč je veľmi komplexný systém zvukov, ktorý vytvára hlasový trakt, pričom najväčšiu úlohu pri tvorbe majú hlasivky. Hlasivky sú schopné vytvárať veľmi zložité zvuky, ktoré predstavujú informácie v podobe zvukových signálov, zakódovaných v určitom jazyku, ktoré v konečnom dôsledku predstavujú ľudskú reč. [1, 2]

1.2.1 Hlasový trakt

Hlasový trakt predstavujú jednotlivé rečové orgány, ktoré môžeme rozdeliť do troch skupín:

1. Dychové ústrojenstvo
2. Hlasové ústrojenstvo
3. Artikulačné ústrojenstvo

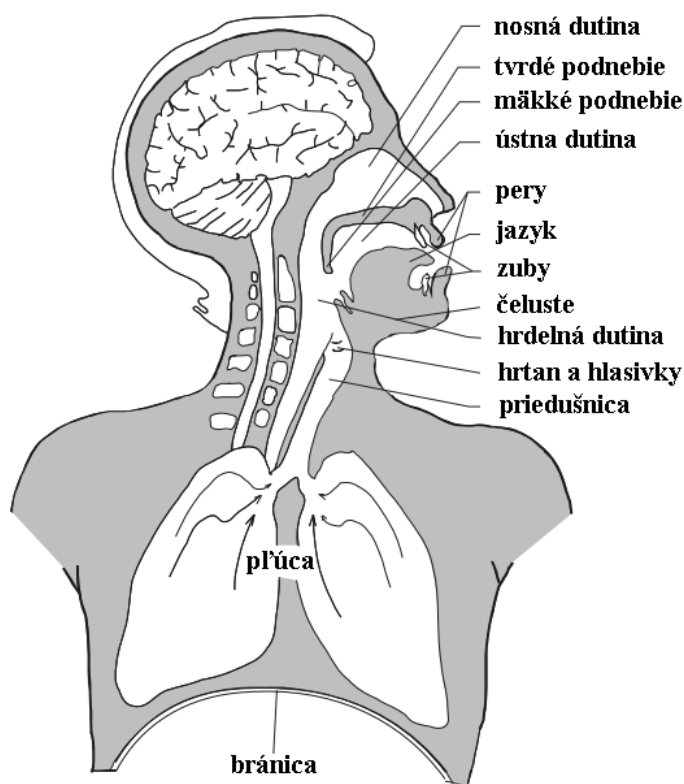
Dychové ústrojenstvo predstavuje základný zdroj energie pre tvorbu reči. Súčasťou tohto ústrojenstva sú pľúca a svaly späť s nimi, bránica (pozri obr. 1.4).

Hlasové ústrojenstvo je uložené v hrtane a najdôležitejšou časťou sú hlasivky. Hlasivky predstavujú dve slizničné riasy, medzi ktorými je priestor nazývaný hlasivková štrbina.

Artikulačné ústrojenstvo sa skladá z nadhrtanových dutín (hrdelná, ústna a nosná dutina) a artikulačných orgánov (jazyk, pery, zuby a mäkké podnebie), pozri obr. 1.4. [1]

Základom pre tvorbu reči je vydychovaný prúd vzduchu vznikajúci v pľúcach za pomoci tlaku bránice. Vydychovaný vzduch ďalej prechádza priedušnicou, hrtanom a hltanom a vychádza ústnou dutinou, resp. nosnou dutinou. Vydychovaný vzduch môžeme v teoretickej

1 EMOČNÉ STAVY A ICH FYZIOLOGICKÝ PREJAV



Obrázok 1.4 Hlasový trakt. [1]

rovine považovať za nosný signál, pričom je tvarovaný jednotlivými časťami hlasového traktu ako napríklad hlasivky. Pokiaľ človek mlčí, tak je hlasivková štrbina odkrytá, aby mohol cez ňu prechádzať vzduch. Naopak pri hovorení sú hlasivky stiahnuté a pod tlakom vydychovaného vzduchu sa prehýbajú a kmitajú (pozri obrázok 1.5). Kmitaním hlasiviek vzniká vzduchová vlna, ktorá predstavuje zvukový signál. Vlastnosti tohto signálu závisia od dĺžky, sily a svalového napätia hlasiviek. Tento zvukový signál predstavuje základný tón reči. [1, 2]

1.2.2 Základný tón reči

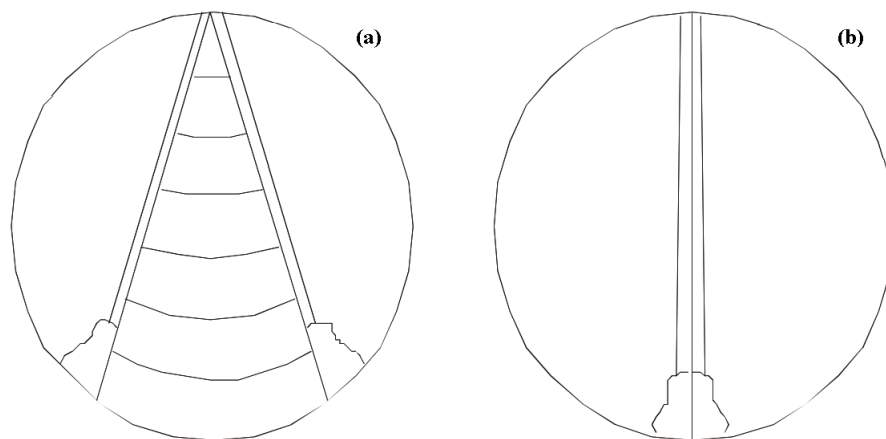
Je to periodický signál, ktorého frekvencia závisí od frekvencie kmitania hlasiviek. Túto frekvenciu označujeme ako F_0 . Perióda signálu základného tónu je označovaná ako T_0 a je vlastne prevrátenou hodnotou frekvencie tohto signálu $T_0 = 1/F_0$. [2]

Pre dospelých ľudí sa základný tón reči typicky pohybuje v rozsahu od 80 do 400 Hz, ale môže sa meniť v rozsahu už od 60 až do 3100 Hz. U žien je v priemere dvakrát vyšší ako u mužov, u detí môže dosahovať hodnotu až 600 Hz. Frekvencia závisí od dĺžky hlasiviek a je nepriamo úmerná. Dĺžka hlasiviek sa pohybuje v rozmedzí približne od 12 až do 18 mm, pre ženy typicky 13 mm a pre mužov 15 mm. [1, 2]

Frekvencia základného tónu nie je konštantná, ale sa v priebehu hlasového prejavu rýchlo mení. Túto zmenu zaznamenáva funkcia jitter a uvádza sa v percentách. Pri neutrálnom hlasovom prejave má jitter hodnotu 1 percenta. Ďalším parametrom, ktorý je pri základnom tóne dôležitý je shimmer. Tento parameter udáva kolísanie amplitúdy hlasového prejavu (hodnota sa

1 EMOČNÉ STAVY A ICH FYZIOLOGICKÝ PREJAV

udáva v dB). Pre normálneho poslucháča je rozoznatelná hodnota zmeny frekvencie od 2 percent vyššie a zmeny amplitúdy od 1 dB.



Obrázok 1.5 (a) hlasivky s odkrytou hlasivkovou štrbinou, (b) so zakrytou hlasivkovou štrbinou. [1]

1.2.3 Vznik formantov

Hlas pozostávajúci len zo základného tónu by nebol takmer vôbec badateľný. Vďaka rezonanciám hrtanovej, nosnej a ústnej dutiny sa mení farba hlasu. Spomenuté dutiny dokážu zosilňovať základný tón. Tieto tóny, vznikajúce v dutinách, nazývame formanty, pričom tón vytváraný v hrtanovej dutine nazývame vedľajším formantom. Hlavný formant, vznikajúci v ústnej dutine, môže mať frekvenčný rozsah od 175 až do 3600 Hz. Zmeny tvarov a polohy pohyblivých artikulačných orgánov umožňujú meniť tvary a rozmery nad hrtanových dutín, čím dochádza ku zmene hlavného formantu alebo vzniku rôznych vlastných zvukov reči (napríklad šum). Zvukový signál, vytvorený na hlasivkách, je zosilnený hlavne v oblasti formantových frekvencií. [1, 2]

Vďaka veľkému frekvenčnému rozsahu hlavného formantu, čo je dôvodom rôzneho zloženia ľudského hlasu, môžeme mať rôzne samohlásky. Hodnoty prvých troch formantov sú znázornené na obrázku 1.6 vo forme tabuľky. [2]

1.2.4 Varianty reči

Existuje viacero variant reči, oproti normálnemu rečovému prejavu človeka, ktorý sa nachádza v neutrálnom stave. Sú to varianty rôznych intenzít a frekvenčných rozsahov, spev, šepot a krik.

Spev je variant reči, pri ktorej dochádza k zmene trvania rečových jednotiek, pričom je snaha o dosiahnutie určitého rytmu. Väčšinou sa mení dĺžka vyslovovania samohlások. Pri speve sa dodáva množstvo energie na 3. a 4. formant, pričom pri normálnej reči k tomu nedochádza. Dochádza hlavne k zmene artikulácie samohlások, čo má za následok vznik tzv. spevového formantu okolo frekvencie 2 800 Hz. [4]

Ďalším, tentoraz netypickým, variantom reči je šepot. Pri šeptaní nedochádza k uzavretiu hlasiviek a zvuk vzniká len trením pri prechádzaní vzduchu cez hlasivky, a preto úplne zanikajú znelé tóny a sú nahradené neznelými. Počas šepotu klesá hladina intenzity z bežnej hladiny 40 –

1 EMOČNÉ STAVY A ICH FYZIOLOGICKÝ PREJAV

60 dB na hladinu 20 – 30 dB. Šepot sa hlavne používa, ak nechceme niekoho vyrušovať alebo ak nechceme, aby nás niekto iný počul. [4]

samohláska	F_1 [Hz]	F_2 [Hz]	F_3 [Hz]
u	300 – 500	600 – 1000	2400 – 2900
o	500 – 700	900 – 1200	2500 – 3000
a	750 – 1100	1100 – 1500	2500 – 3000
e	500 – 700	1500 – 2000	2500 – 3000
i	300 – 500	2000 – 3000	2600 – 3000

Obrázok 1.6 Hodnoty prvých troch formantov pre samohlásky. [1]

Posledným variant, ktorý spomeniem, je krik. Jedná sa o extrémnejší variant reči, pretože intenzita prejavu stúpa z hladiny 40 – 60 dB až na 80 dB. Pri kričaní dochádza k zmene napätia hlasiviek, ale je aj zvýšený tlak vydychovaného vzduchu, je vydychovaný rýchlejšie. Tým je skrátená doba otvárania hlasiviek. Pri kričaní dochádza k zvýrazneniu samohlások nielen zvýšením intenzity, ale aj dĺžkou trvania. Tým, že pri kriku dochádza k väčšiemu otváraní úst, tým dochádza k zmene prvého formantu a frekvenčného spektra. [4]

1.3 Vlastnosti reči

Parametre signálu pri spracovaní reči sa nazývajú vlastnosti. V oblasti spracovania reči, je veľa vlastností, ktoré sú využívané pri rozpoznávaní emócií. Táto časť práce opisuje parametre, ktoré sú používané pri rozpoznávaní emócií z reči. Tieto vlastnosti môžu byť rozdelené do viacerých kategórií. Vo viacerých prípadoch sú používané prozodické, spektrálne, hlasovo kvalitatívne a lingvistické vlastnosti. [9]

1.3.1 Prozodické vlastnosti

Človek rozpoznáva prozodické charakteristiky ako rozsah, hlasitosť, tempo rozprávania, pauzu a rytmus. Tieto charakteristiky reči sú rozpoznávané človekom za pomoci hraníc vo vetách a v texte ako čiarka. Takisto aj nálada s ako je veta vyslovená alebo napísaná je dôležitá. Je úplne rozdielne pokiaľ mám vety „Bude to v nedeľu?“ a „Bude to v nedeľu!“ . Zachytiť tu iróniu vo vyslovenej vete pri komunikácii dvoch ľudí je tiež niekedy problém. V rečových signáloch nemáme žiadny ekvivalent, ale existujú nejaké vlastnosti, ktoré korelujú s týmito charakteristikami. Základný tón (F_0) môže opisovať vnímaný rozsah a krátkodobá energická vlastnosť koreluje s hlasitosťou. Prozodické vlastnosti v spracovaní reči sú vlastnosti základného tónu, vlastnosti energie signálu reči a vlastnosti trvania. [9]

Vlastnosti základného tónu opisujú časovú zmenu základného tónu počas jednej frázy, slova, sylaby alebo iných kratších intervalov. Tieto intervaly sú nazývané segmenty. Základný tón opisuje frekvenciu hlasiviek a je spojený s podráždením. Preto je dobre spomenúť, že základný tón je vyňať len z hlasových častí reči, pretože nehlasové časti hlasu sú vytvárané

1 EMOČNÉ STAVY A ICH FYZIOLOGICKÝ PREJAV

roztvorenými hlasivkami. Základná funkcia tohto tónu je definovaná kontúrou počas intervalu. Hodnoty, ktoré sú najviac brané do úvahy sú priemer, štandardná odchýlka, medián, rozsah a minimum s maximom. [9]

Vlastnosti energia rečového signálu sú takmer rovnaké ako tie od základného tónu a opisujú zmenu energie signálu počas intervalu. Funkcionály sú používané na kontúru hodnôt signálovej energie v krátkom časom intervale.

Vlastnosti trvania sú ovplyvňované štýlom hlasového prejavu. Najviac je to merané v časovej doméne ako fonémy, slová alebo celé frázy. Tieto vlastnosti sú často merané ako trvanie hlasových a nehlasových častí reči. Ako som spomenul, štýl hlasového prejavu ovplyvňuje trvanie. To znamená, že pauzy v hlasovom prejave priamo menia vlastnosti trvania. Priemer, medián, štandardná odchýlka a podobne, sú vlastnosti trvania rovnako ako v prípade základného tónu a vlastnosti energie. Tempo reči je najviac používaná vlastnosť trvania, ktorá je definovaná ako pomer aktuálne skúmanej dĺžky a predpokladanej dĺžky. Pomer dĺžky pauzy počas reči, najdlhšej trvanie pauzy a počet najdlhších páuz sú niektoré z vlastností trvania ktoré majú súvislosť s pauzami v rečovom prejave. Pozícia základného tónu a energia extrémov sú tiež vlastnosti trvania. [9]

1.3.2 Spektrálne vlastnosti

Spektrálne vlastnosti opisujú charakteristiky rečového signálu vo frekvenčnej doméne. Prozodické vlastnosti zahŕňajú základný tón, energiu a trvanie. Na rozdiel od tohto spektrálne vlastnosti opisujú harmonické frekvencie a formanty. Harmonické frekvencie môžu byť nájdené ako násobky základného tónu v spektre. Frekvencia a amplitúda, ktorá definuje každú harmonickú frekvenciu, sú často používané pre rozpoznávanie emócií. Formanty sú vznikajú rezonanciou v dutinách hlasového traktu a môžu byť opísané frekvenciou, amplitúdou a šírkou pásma. Zvuky hlasu obsahujú od 3 do 4 formantov (malá šírka pásma, 8 kHz), pozri rovnica 1.1, kde f_s je vzorkovacia frekvencia. Existuje veľa štúdií rozpoznávania emócií, ktoré používajú formanty ako parametre. [9]

$$n_{formant} = \frac{f_s [kHz]}{2} - 1 \quad (1.1)$$

Pre rozpoznávanie emócií môže byť tiež používaný energický stĺpec, ktorý vlastne predstavuje pomer spektrálnej energie nad a pod určitou frekvenciou, napríklad 800 Hz. Dodatočná vlastnosť je väčšinou používaná pri rozpoznávaní reči. Aj keď sa nejedná o spektrálnu doménu, Mel-frekvenčné spektrálne koeficienty (MFCC) sú zoskupované ako spektrálne vlastnosti. Tieto parametre boli vyvinuté pre zisťovanie obsahu, ale neskôr sa zistilo, že majú veľký prínos ako sekundárne informácie z reči. MFCC bolo úspešne používané v rozpoznávaní reči, ktoré bolo zamerané na „Ako to bolo povedané“, špeciálne na rozpoznávanie emócií. Tieto vlastnosti môžu byť priemerované nad celým segmentom alebo na úrovni rámcov (Dlhodobé priemerné spektrum (LTAS z anglického výrazu) alebo Krátkodobé priemerné spektrum). MFCC môže byť nahradené inou dobre známou vlastnosťou z rozpoznávania reči ako lineárna prediktívne cepstrálne koeficienty (LPCC). Na rozpoznávanie reči sú tiež používané lineárne spektrálne frekvenčné (LSF) parametre, ktoré sú alternatívou k lineárnym predikčným parametrom. [9]

1 EMOČNÉ STAVY A ICH FYZIOLOGICKÝ PREJAV

1.3.3 Vlastnosti hlasovej kvality

Kvalita hlasu je charakteristická pre vokálny štýl, ako napríklad modálny (neutrálny) hlas, bez zvučný, šepkajúci, vŕzgajúci, krutý alebo falsetto hlas. Jeden z výskumov hovorí: „Napriek tomu, že parametre základnej frekvencie (viazané na rozsah) sú nepochybne dôležité vo hlasovom vyjadrovaní emócie, aj tak kľúčom k hlasovej diferenciácii diskkrétnej emócie vyzerá byť hlasová kvalita“. To znamená, že hlasová kvalita môže hrať významnú úlohu v rozpoznávaní emócií. Jedným z problémov, ktoré nastávajú pri riešení hlasovej kvality, je určenie podráždenia hlasového traktu. Existuje niekoľko návrhov na automatické inverzné filtrovanie, avšak tieto metódy merania sú náchylné na šum. [9]

Okrem iného inverzné filtrovanie je používané pre merania hlasovej kvality riadené pomocou spektra rečového signálu. Jitter je odchýlka druhu cyklus-cyklus základného tónu hlasu. Relatívny jitter je v vzťahu s priemernou periódou, ktorá znázornená vo vzťahu 1.2 nižšie.

$$Jitter_{REL} = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i+1}|}{\frac{1}{N} \sum_{i=1}^N T_i} \quad (1.2)$$

Ďalším dobre známym parametrom hlasovej kvality je shimmer, ktorý odráža odchýlku typu maximum-maximum jednotlivých amplitúd a uvádza sa v decibeloch. Vzťah je uvedený nižšie pod číslom 1.3.

$$Shimmer(dB) = \frac{1}{N-1} \sum_{i=1}^{N-1} 20 \log \left(\frac{A_{i+1}}{A_i} \right) \quad (1.3)$$

Podobne ako pri jitteri, tak aj relatívny shimmer je rozdelený pomocou priemernej amplitúdy. Parameter harmonická-ku-šumu (HNR z anglického výrazu) patrí do skupiny vlastností hlasovej kvality a meria sa v stupňoch periodicity. Mali by sme si poznamenať, že pre rozpoznávanie reči sú často používané lingvistické vlastnosti. [9]

1.3.4 Prozodické vlastnosti českého jazyka

Český jazyk je jazykom, kde každú vetu si je možné predstaviť ako jasnú melodickú a ucelenú jednotku. Podľa jedného zdroja z 1994 je hlavným prvkom vytvárajúcim vetu melódia. Samozrejme rovnako aj intenzita a časové rozdelenie prispievajú k organizovaniu tejto jednotky. Podľa štúdií z 1951 a 1957 sú zamerané na rozpoznanie všetkých druhov organizácií a sú podstatou je sluchová analýza. [4]

Český jazyk má prízvuk zviazaný s prvou slabikou rytmického celku. Niektoré české nárečia môžu mať prízvuk na predposlednej slabike. Tento prízvuk je avšak používaný len v pohraničných oblastiach, kde je český jazyk ovplyvňovaný inými jazykmi. V českom jazyku je hlavne dynamický charakter prízvuku a to preto, že sa na jeho vytváraní podieľa hlavne intenzita. Prízvuk v slove nemá taký význam ako napríklad v ruskom jazyku. V ruskom jazyku to isté slovo s rôznymi prízvukmi znamená niečo iné. Prízvuk v češtine nie je tvorený len intenzitou, ale aj zmenou výšky a dĺžky trvania. Pokiaľ je prostredie v ktorom subjekt rozpráva ideálne, teda nie je subjekt rušený inými zvukmi alebo okolím, tak je použitá vyššia intenzita a základnú frekvenciu. Tempo je avšak na druhej strane pomalšie. Pokiaľ nie je prostredie,

1 EMOČNÉ STAVY A ICH FYZIOLOGICKÝ PREJAV

v ktorom subjekt rozpráva, ideálne, sa preukázalo, že dostatočne postačuje spomalenie tempa a vytvorenie dojmu prízvuchej slabiky, aj napriek tomu, že nebolo možné zvýšiť intenzitu a základnú frekvenciu. [4]

Hlavná časť hovorovej reči zvyčajne je umiestnená až na konci rečového prejavu. Touto časťou je spojené umiestnenie prízvuku vo vete. Prízvuk vo vete je realizovaný pomocou nízkeho tónu, ktorý ľahko rastie ku koncu slabiky. [4]

Tempo sa z veľkej časti podieľa na vnímaní rytmu rečového prejavu. Zmeny, vykonané na tempe, predstavujú zvukovú zmenu, ktorá v českom jazyku predstavuje citové zaujatie rozprávajúceho subjektu. Každý rozprávajúci človek má svoje základné tempo, resp. rýchlosť rečového prejavu. V súčasnom českom jazyku sa hodne rozšírilo koncové predlžovanie. Uvediem príklad: ak je v poslednej slabike slova dlhá samohláska, môže mať táto slabika trvanie až o 370 percent dlhšie ako priemerný krátky vokál. Ku príkladu, trvanie krátkeho vokálu môže dosiahnuť hodnoty až 145 percent priemeru. [4]

1.4 Objektívne metódy pre určovanie emócií

Existuje veľa objektívnych metód na meranie emócií a stresu. Elektro-fyziologické metódy a laboratórne hodnotenia hladiny stresových hormónov sú dosť bežne v neurológii a psychológii. Elektro-fyziologické metódy zahŕňajú neinvazívne merania za pomoci elektro-kardiografie (ECG), elektro-encefalografie (EEG), elektro-okulografiu (EOG), záznam činnosti jednotlivých neurónov za pomoci magneto-ecefalografie (MEG), pozičná emisná tomografia (PET), funkčnú magnetickú rezonanciu (fMRI), optickú snímanie ako napríklad blízko infračervené snímkovanie (fNIR) a ďalšie invazívne elektro-fyziologické metódy ako napríklad elektro-kortikografia (ECoG). Taktiež jednoduché meranie odporu kože je celkom užitočné v hodnotení stresu. I keď metódy merania ako napríklad MEG, PET, fMRI a fNIR sú celkom finančne náročné, závisia na metabolických procesoch a vyhodnotenie týchto procesov trvá dlho, takže sú menej vhodné pre rýchlu komunikáciu. V súčasnosti neinvazívne a invazívne metódy ako ECG, EEG, ECoG a záznam jedného neurónu sú jediné metódy, ktoré sú relatívne jednoduché a cenovo dostupné, pričom zaručujú dočasné riešenie s vysokou úspešnosťou. [10]

Ďalšou z objektívnych metód na analýzu a získanie emócie ľudského subjektu je možné použiť neurálny klasifikátor emócií. Tento klasifikátor pracuje na princípe fungovania ľudského mozgu a teda používa metódu analýzy, ktorá sa nazýva neurónová sieť. Viac detailov je možné nájsť v tomto literatúre [11, 12]. Jedná sa však o pomalú vyhodnocovaciu metódu, ktorá má viacej krokov a jej výsledky závisia od viacerých parametrov, resp. v prípade zlého nastavenia môže byť menej presne ako predchádzajúce metódy.

2 VYTVÁRANIE EMOČNEJ DATABÁZY

2 VYTVÁRANIE EMOČNEJ DATABÁZY

V tejto kapitole sa budem venovať, ako už názov naznačuje, tomu ako som vytváral databázu emočných vzoriek. Spomením aké požiadavky boli kladené na databázu na základe zadania. Ďalej budem aj popisovať ako som si pomenoval vzorky databázy tak, aby boli jednoznačne pomenované. Budem v krátkosti popisovať aké zdroje som používal pri vytváraní tejto databázy a spomením aj aké parametre by mali mať zvukové nahrávky. A nakoniec spomením program za pomoci ktorého som strihal jednotlivé vzorky.

2.1 Požiadavky na databázu

V zadaní 2. bodu práce je dané, že mám vytvoriť emočnú databázu. Zvyšok zadania 2. bodu mi bol mojím vedúcim práce zdelaný v ústnom podaní. Mal som vytvoriť databázu, ktorá bude obsahovať zvukové vzorky s nehranými reálnymi emóciami. Teda emočné vzorky nemajú mať pôvod v hraných emóciách. To znamená, že som nemohol čerpať z kinematografie. Musel som používať zdroje, resp. zvukové záznamy, ktoré obsahovali reálne emócie. Použil som zvukové záznamy relácií z archívov rádií a televízie. Niektoré záznamy relácií neboli umiestnené na oficiálnych stránkach, a preto som použil aj záznamy, ktoré boli umiestnené na video zdieľajúcom portáli, YouTube .

Ďalej mali jednotlivé zvukové vzorky spĺňať nasledujúce požiadavky:

- zvuková nahrávka ma byť dlhá od 1 až do 6 sekúnd,
- zvuková nahrávka by nemala obsahovať okolitý ruch,
- zvuková nahrávka musí obsahovať hovorené slovo,
- zvuková nahrávka musí byť unikátne pomenovaná.

To znamená, že každá vzorka musí obsahovať aspoň časť vety, resp. celú vetu. Osoba, ktorej hlasový prejav bude na nahrávke by mal byť zreteľný a nemal by byť rušený iným hlasom alebo ruchom na pozadí. Vzorka musí byť jednoznačne pomenovaná tak, aby bolo jasné o aký emočný stav ide, resp. koľká vzorka je v poradí pre danú emóciu a o aké pohlavie hovoriacej osoby sa jedná.

2.2 Pomenovanie nahrávok emočnej databázy

Jednou z požiadaviek, ako bolo uvedené vyššie v podkapitole, bolo jednoznačne pomenovať vzorku databázy tak, aby pri hodnotení vzoriek nedošlo k zámene, resp. iným problémom. Názov mal obsahovať druh emočného stavu a pohlavie osoby, ktorej emočný stav bol na vzorke zaznamenaný. Aby bola vzorka jednoznačne rozpoznateľná, tak bolo nutné použiť poradové číslo pre vzorku danej emócie. Teda názov vzorky je zložený s týchto vecí v danom poradí: emočný stav osoby, poradové číslo pre daný emočný stav a na poslednom mieste je pohlavie osoby, ktorej emočný stav je zaznamenaný. Názov má teda celkovú dĺžku päť písmen, pozri obrázok 2.1.

Na prvom mieste názvu je jedným písmenom označené pohlavie. Použil som anglické výrazy, z ktorých som použil počiatočné písmena pre jednotlivé emočné stavy:

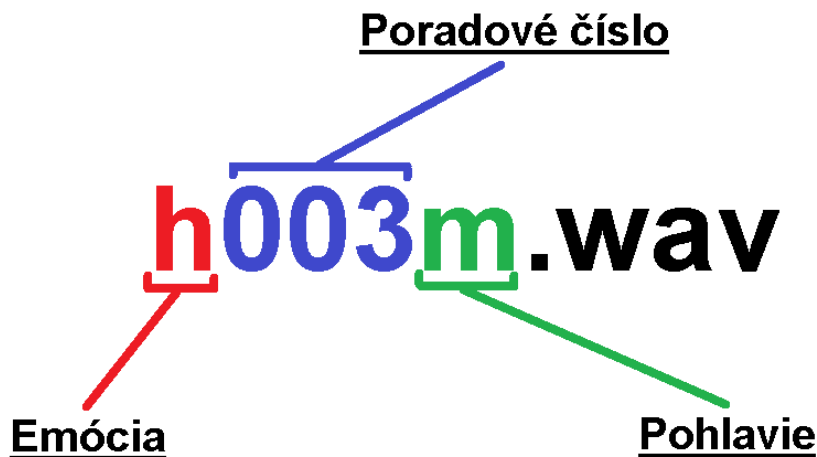
2 VYTVÁRANIE EMOČNEJ DATABÁZY

- n = normal / neutral state, neutrálny / normálny emočný stav,
- a = anger, hnev,
- f = fear, strach,
- h = happiness, radosť alebo šťastie,
- s = sadness, smútok.

Na druhom až štvrtom mieste je teda poradové číslo pre daný emočný stav, nie pre všetky emočné stavy. Na poslednom mieste názvu je pohlavie osoby, ktorej emócia bola zachytená. Opäť som použil počiatočné písmeno z anglického výrazu pre obe pohlavia:

- m = male, mužské pohlavie – muž,
- f = female, ženské pohlavie – žena.

Príklad názvu je uvedený nižšie na obrázku 2.1, kde bola použitý ako emočný stav šťastie, pohlavie muž a poradové číslo je jasné.



Obrázok 2.1 Zloženie názvu pre jednotlivé zvukové nahrávky.

2.3 Parametre nahrávok emočnej databázy

Aby zvukové nahrávky mali určitú kvalitu a zároveň nezaberali veľa miesta na disku, tak musia mať určité parametre. Emočnú databázu som primárne vytváral pre doktoranda Pavla Partilu, takže na základe ústnej dohody s ním som si stanovil parametre zvukových nahrávok. Zvukové nahrávky mali byť nekomprimované, teda som použil bezstratový zvukový formát wave alebo Waveform Audio File Format. Aj preto sú parametre nahrávok nasledujúce:

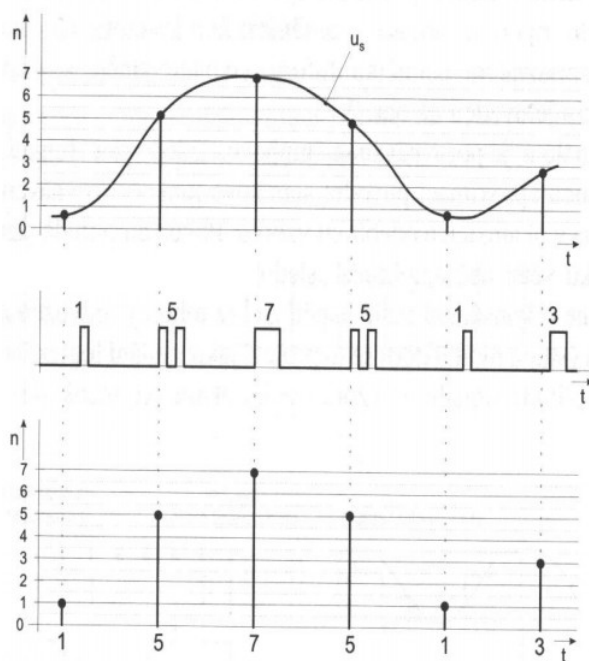
- kódok PCM s16 le (s16l),
- z toho vyplýva bit rate 16 bit,
- jeden kanál = mono,
- vzorkovacia frekvencia 16 kHz.

Všetky tieto parametre dávajú dokopy celkový dátový tok 256 kb/s. Vzhľadom ku kvalite zdrojových nahrávok, z ktorých som strihal bola vzorkovacia frekvencia 16 kHz dostačujúca.

2 VYTVÁRANIE EMOČNEJ DATABÁZY

Frekvencia vyššia od zvolenej by zvyšovala kvalitu nahrávky, avšak na úkor veľkosti nahrávky. Na druhu stranu nižšia vzorkovacia frekvencia by aj výrazne znížila kvalitu nahrávky. Použitie jedného kanálu, teda mono záznamu tiež napomohlo k zníženiu veľkosti nahrávky a to presne o polovicu, keďže zdrojové nahrávky boli stereo.

Vzhľadom k tomu, že som si zvolil wave ako výstupný audio formát, tak s ním je úzko spätá modulácia PCM, skratka s anglického výrazu Pulse Code Modulation. Podstatou tejto modulácie sú tri operácie: vzorkovanie, kvantovanie a kódovanie. Spôsob akým funguje táto modulácia je znázornený na obrázku 2.2 nižšie. Priebeh signálu je najskôr navzorkovaný. Aktuálne hodnoty, v tom danom momente, sú priradené k určitej kvantizačnej úrovni n . Na obrázku nižšie je použitá 3 bitová PCM modulácia. To znamená, že ku každej úrovni je priradená určitá kódová kombinácia. V tomto prípade je možné jeden stav zapísať pomocou troch bitov do kódovej kombinácie. Teda 2 na 3 stavov, čo celkovo predstavuje 8 stavov (od 0 do 7), resp. kvantizačných úrovní. V mojom prípade som použil 16 bitovú PCM, ktorá dokáže aktuálnu hodnotu priradiť až ku 65 536 kvantizačným stavom. Táto kódová kombinácia je výstupným produktom PCM modulácie a je určená k prenosu pomocou kanálu. Keďže pri prenose dát kanálom dochádza k rušeniu a tým aj k skresľovaniu tohto digitálneho signálu, tak je nutné tento signál na prijímacej strane regenerovať a až následne dekódovať. [8]



Obrázok 2.2 Princíp akým funguje PCM modulácia a demodulácia. V hornej a strednej časti je zobrazené vzorkovanie, kvantovanie a kódovanie, teda proces modulácie a v spodnej časti obrázka je demodulácia. [8]

Presnosť aktuálnych hodnôt jednotlivých vzoriek závisí na kvantizačnom kroku. Inak povedané, závisí to na počte bitov, resp. počte kvantizačných úrovní, teda čím viac úrovní, tým je krok menší a tým presnejšie je možné modulovať pôvodný signál. Samozrejme presnosť modulácie, resp. kvalita digitalizovanej nahrávky závisí aj na vzorkovacej frekvencii. V mojom

2 VYTVÁRANIE EMOČNEJ DATABÁZY

prípade som použil vzorkovaciu frekvenciu 16 kHz, čo znamená, že výsledná digitalizovaná nahrávka má 16 000 hodnôt za 1 sekundu a každá jedna hodnota je zapísaná do 16 bitov. [8]

Veľkou nevýhodou PCM modulácie je relatívne veľká šírka potrebného frekvenčného pásma. Naopak výhodou PCM je odolnosť voči rušivým napätiam, pokiaľ nie je presiahnutá určitá hodnota. [8]

2.4 Zdrojové nahrávky

Vzhľadom k tomu, že malo ísť o databázu, ktorá bude zhotovená z nahrávok obsahujúcich reálne nehrané emócie, tak som niektoré nahrávky stiahol z internetových archívov, niektoré som nahral z internetového streamu a niektoré som si stiahol z Youtube.

Niektoré zdrojové nahrávky som získal nahrávaním zo streamu pomocou VLC ako sú:

- relácie z rádia Frekvence 1.

Konkrétne sa jednalo o relácie Dámsky klub, Ráno s Frekvencií 1 a Ruda z Ostravy. Použil som stream zo serveru play.cz, ktorý je poskytovateľom streamingovej služby pre väčšinu českých rádii. Konkrétne to bol stream <http://icecast4.play.cz:8000/frekvence1-128.mp3>. Podľa názvu linku je možné vidieť, že sa jednalo o mp3 stream o celkovom dátovom toku 128 kb/s. Okrem toho nahrávka mala vzorkovaciu frekvenciu 44 kHz, takže kvalita zdrojovej nahrávky bola relatívne dobrá. Zvolil som takýto spôsob z toho dôvodu, že rádio Frekvence 1 má archív, z ktorého nie je možné priamo sťahovať záznamy relácii na rozdiel od ďalšieho rádia, ktoré spomeniem nižšie.

Ďalšie zdrojové nahrávky som si stiahol z internetového archívu:

- relácia z Fajn rádia.

Jednalo sa o reláciu Pomsta je sladká na Fajn rádiu. Opäť ako v predchádzajúcom prípade bola kvalita nahrávky rovnaká, teda vzorkovacia frekvencia 44 kHz a celkový tok 128 kb/s. Archív tohto rádia nájdete na adrese <http://www.fajnradio.cz/fajn-radio/archiv-podcast/pomsta-je-sladka-strana-0/>.

Poslednou možnosťou ako som získal nahrávky k spracovaniu na emočné vzorky bolo stiahnutie si týchto nahrávok zo serveru YouTube. Boli to nasledovné nahrávky z:

- relácia z Fajn rádia,
- relácia z Českej Televízie.

YouTube je vlastne databáza krátkych videí, ktoré sú uložené vo formáte flashového videa, teda majú koncovku .flv. Okrem toho videa sa priamo nedejú sťahovať z YouTube, resp. na stránkach YouTube nie sú priame odkazy na stiahnutie. Zvažoval som dve možnosti ako získať relácie z YouTube:

- za pomoci voľne dostupného pluginu pre prehliadač,
- pomocou stránky <http://www.youtube-mp3.org/>.

Keďže som potreboval len zvukovú stopu záznamu bez videa, v prípade relácie z ČT, tak som si zvolil druhú možnosť. V prípade prvej možnosti by som si stiahol záznam vo formáte

2 VYTVÁRANIE EMOČNEJ DATABÁZY

flashového videa a musel by som to konvertovať do mp3. V druhom prípade to stránka urobí za mňa a priamo si môžem stiahnuť len zvukový záznam vo formáte mp3 a o rovnakej kvalite ako v prípade predchádzajúcich možností, teda 44 kHz vzorkovacia frekvencia a 128 kb/s celkový dátový tok. Z Fajn rádia sa opäť jednalo o reláciu Pomsta je sladká. V prípade relácie Českej televízie išlo o Poštu pro tebe.

2.5 Spracovanie zdrojových záznamov

Na základe ústneho dohovoru s doktorandom Pavlom Partilom, pre ktorého som mal zhotoviť databázu emočných vzoriek, som si stanovil, že vytvorím databázu zloženú z piatich emočných stavov. Konkrétne to mali byť neutrálny stav, hnev, šťastie, smútok a strach. Získavanie vzoriek bolo veľmi časovo náročné a práce, pretože niektoré záznamy obsahovali veľa častí, ktoré sa nedali spracovať. Ako napríklad v prípade záznamov z rádia Frekvence 1, veľké časti tvorila hudba, resp. často sa stávalo, že časť záznamu obsahovala okrem hlasu na pozadí aj hudbu, čo sa považovalo za rušivý faktor a nebolo to vhodné na spracovanie.

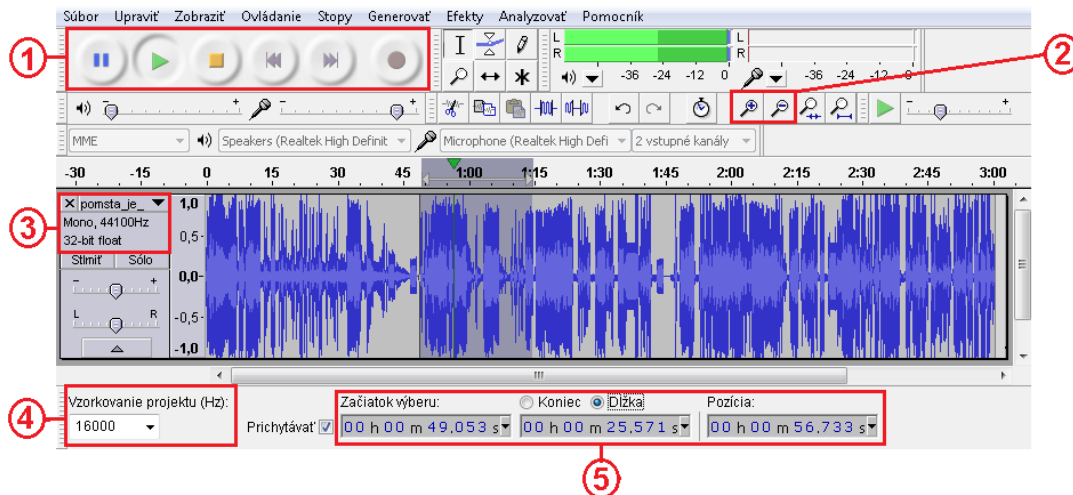
V prípade všetkých v predchádzajúcej kapitole spomenutých relácií bolo možné získať zo všetkých záznamov vzorky neutrálnej emócie. Vzorky hnevu bolo možné získať hlavne zo záznamov relácie Pomsta je sladká, ale nejaké vzorky sa dali získať aj z relácie Pošta pro tebe. V prípade šťastia sa dali vzorky získať z relácií Pošta pro tebe a z relácií rádia Frekvence 1. V prípade smútku bolo zložité a časovo náročné získať vzorky, ale nejaké som vystrihol z relácie Pomsta je sladká a z relácií rádia Frekvence 1. Avšak najväčším zdrojom vzoriek smútku bola relácia Pošta pro tebe. Chcel som získať aj vzorky strachu, ale v tomto prípade to bolo veľmi náročné a celkovo som našiel len málo vzoriek. Takže som sa rozhodol, samozrejme so súhlasom doktoranda, že databázu obmedzíme na štyri emočné stavy. Samozrejme doktorand Partila vyžadoval po mne, čo najväčšiu databázu, avšak sa mi podarilo zhromaždiť len 50 vzoriek pre 3 emočné stavy a pre neutrálny stav som vystrihol takmer 150 vzoriek. Avšak finálna databáza, ktorá bola hodnotená obsahovala len po 50 vzoriek z každej emócie, aby to bolo rovnomerné.

Aby som sa vrátil k spracovaniu zdrojových nahrávok. Tieto získané záznamy som potreboval nejakým spôsobom spracovať. Na internete nájdete kopu platených programov, ktoré dokážu strihať zvukové záznamy. Ale ja som si chcel najisto nejaký bezplatný program, ktorý by bol jednoduchý a prehľadný.

Všetky predpoklady spĺňal program Audacity, ktorý je priamo zameraný len na editovanie a vytváranie zvukových záznamov. Po inštalácii tohto programu som zistil, že v základe nepodporuje zvukový formát mp3. Preto som hľadal na oficiálnej stránke riešenie. Bolo nutné podľa inštrukcií na oficiálnej stránke doinštalovať Lame Encode knižnicu a program bez problémov podporoval formát mp3. Užívateľské rozhranie programu je jednoduché a relatívne intuitívne, je zobrazené nižšie na obrázku 2.3. Niektoré funkcie boli od začiatku jasne, funkčnosť niektorých som si musel zistiť na internete. Výstupné parametre všetkých vzoriek boli také, aké som spomínal v podkapitole 2.3 Parametre nahrávok emočnej databázy.

Aby som dostal finálnu nahrávku, ktorá mala mať tieto parametre, tak som najprv musel zmeniť odstrániť z nahrávky jeden kanál, aby som dostal zo stereu mono. Potom som musel zmeniť vzorkovaciu frekvenciu zo 44 kHz na 16 kHz. A nakoniec pri exportovaní časti vzorky z nahrávky, som túto vzorku uložil ako PCM 16 bit wave súbor. Zdrojové nahrávky mali 32 bit bit rate.

2 VYTVÁRANIE EMOČNEJ DATABÁZY



Obrázok 2.3 Rozhranie programu Audacity. V obrázku sú popísané hlavné časti rozhrania, ktoré som používal počas extrakcie vzoriek zo zdrojových nahrávok. 1 – hlavný panel prehrávania, 2 – približovanie a oddaľovanie, pri zameriavaní sa na určitú časť vzorky, 3 – základné parametre nahrávky, 4 – výstupná vzorkovacia frekvencia, 5 – časový panel označujúci počiatok výberu z nahrávky, celkovú dĺžku výberu a aktuálnu pozíciu kurzora pri prehrávaní záznamu.

3 NÁVRH A REALIZÁCIA HODNOTENIA EMÓCIÍ

3 NÁVRH A REALIZÁCIA HODNOTENIA EMÓCIÍ

V tejto kapitole sa budem zaoberať samotným návrhom a realizáciou nástroja pre subjektívne hodnotenie emočných vzoriek. Samozrejme samotne hodnotenie vzoriek nebudem popisovať, keďže sa jedna v celku o jednoduchý proces. To zistíte, keď si prejdete túto kapitolu. Kapitola je zložená z celkovo piatich častí.

V prvej podkapitole popisujem voľbu metódy, akou by som mohol hodnotiť vytvorenú databázu emočných vzoriek. Po voľbe metódy popisujem kroky, ktoré treba vykonať, aby som mohol vytvoriť nástroj na hodnotenie vzoriek databázy. V ďalšej časti popisujem prípravu webového a databázového servera, aby som mohol realizovať návrh nástroja na hodnotenie. V nasledujúcej časti píšem o návrhu vzhľadu nástroja a z čoho všetkého je zložený. Resp. vysvetľujem štruktúru, ako sú jednotlivé časti medzi sebou prepojené. Vo štvrtej podkapitole opisujem návrh databázy, resp. jej tabuľky. Databáza bola navrhnutá tak, aby boli oddelené informácie o vzorkách od dát hodnotenia jednotlivých vzoriek. V poslednej podkapitole opisujem jednotlivé časti skriptu každej zo stránok, z ktorých je nástroj zložený a vysvetľujem ich funkčnosť.

3.1 Voľba subjektívnej metódy hodnotenia

Pre hodnotenie emócií pomocou subjektívnej metódy určite existuje viac možností. Avšak ja som sa rozhodol medzi dvoma možnosťami. Prvú mi navrhol vedúci pán Vozňák. Evaluácia vzoriek databázy by prebiehala za pomoci telefónnej ústredne. Pričom na ústrední by bolo nahrané hlasové menu a pomocou tlačidiel by prebiehalo hodnotenie jednotlivých vzoriek, resp. načítanie ďalšej vzorky alebo opätovné prehranie aktuálnej vzorky. Ako nápad to nebol zlý, avšak moje skúsenosti s telefónnymi ústredňami sú takmer nulové.

Druhou možnosťou, ktorú som navrhol ja, bolo použitie webovej stránky a prepojenie stránky s MySQL databázou. Vzhľadom k tomu, že s vytváraním webových stránok mám viac skúsenosti ako s nastavovaním telefónnych ústrední, tak som si zvolil druhú možnosť.

Takže k vytvoreniu stránky som potreboval postupovať v nasledujúcich krokoch:

- nainštalovať si LAMP (Linux + Apache + MySQL + PHP)
- vytvoriť vzhľad stránky
- vytvoriť štruktúru nástroja na evaluáciu
- navrhnuť databázu na ukladanie výsledkov hodnotenia
- návrh algoritmov nástroja

V prvom bode sa bude jedna o jednoduchú inštaláciu webového servera s podporou PHP skriptovacieho jazyka a databázového servera. V druhom a tretom kroku sa bude jednať o vytvorenie a prepojenie jednotlivých stránok nástroja s tým, že každá bude mať nejakú funkciu v rámci toho nástroja. Štvrtým krokom bude vytvorenie databázy tak, aby obsahovala informácie o vzorke, ale aj výsledky evaluácie jednotlivých vzoriek. Posledným krokom bude vytvorenie algoritmov v skriptovacom jazyku PHP tak, aby bolo všetko prepojené a fungovalo podľa predstáv.

3 NÁVRH A REALIZÁCIA HODNOTENIA EMÓCIÍ

3.2 Príprava pre návrh a realizáciu nástroja

Príprava servera by mala začať inštaláciou operačného systému Linux. Lenže v mojom prípade inštalácia nebola nutná. Na požiadanie mi bol pridelený virtuálny počítač. K PC som dostal pridelenú IP adresu 158.196.244.198 a prihlasovacie meno a heslo k účtu na virtuálne PC. Takže už mi ostávalo len doinštalovať ostatok, aby mi fungoval server.

K inštalácii webového servera Apache 2 som potreboval do príkazového riadku akurát zadať príkaz: *apt-get install apache2*.

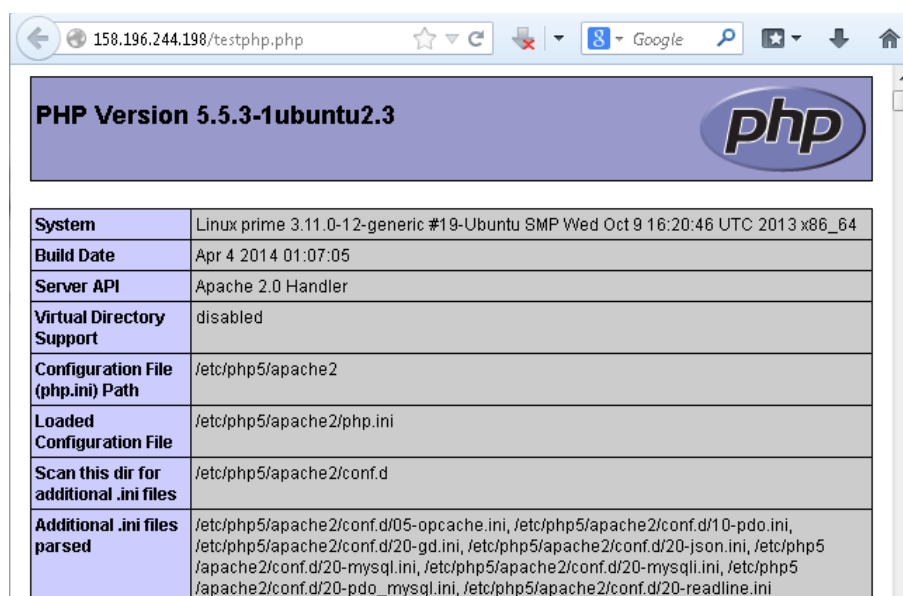
Prebehla inštalácia servera a následne som si otvoril pridelenú IP adresu v prehliadači, aby som si overil, že sa podarilo nainštalovať server správne. V prehliadači mi vypísalo hlášku *IT WORKS!* a samozrejme inštalácia bola úspešná.

Ďalším krokom bola inštalácia knižníc na podporu skriptovacieho jazyka PHP. Na to som potreboval nasledujúci príkaz: *apt-get install php5 libapache2-mod-php5*.

Po ukončení inštalácie PHP verzie 5 bolo nutné reštartovať webový server Apache. Následne, aby som si overil, že inštalácia PHP prebehla v poriadku, tak som musel upraviť súbor *testphp.php* umiestnený v zložke */var/www/*, kde som pridal nasledujúci riadok: *<?php phpinfo(); ?>*. Aby som zistil, že to funguje ako má, tak som si v prehliadači k IP adrese servera dopísal ešte názov súboru, ktorý som upravoval a zobrazila sa stránka s informáciami čo všetko aktuálne server podporuje a aké nastavenia boli použité. Čiastočný výpis spomenutých vecí je na obrázku 3.1 nižšie.

Ďalším krokom bola inštalácia MySQL, teda databázového servera. Potreboval som k tomu príkaz: *apt-get install mysql-server*.

Počas inštalácie mi bola ponuknutá možnosť nastavenia hesla, tak som si nastavil svoje heslo. Ponechanie základného, teda žiadneho hesla, je nebezpečná možnosť. Po ukončení inštalácie som musel upraviť súbor *my.cnf* umiestnený v zložke */etc/mysql/*. Zmenil som riadok *bind address = 127.0.0.1* a prepísal predvolenú adresu na moju IP adresu. Keď by som to neurobil, tak MySQL server by bol dostupný na lokálnom PC a nie v celej sieti, kde je umiestnený server.



System	Linux prime 3.11.0-12-generic #19-Ubuntu SMP Wed Oct 9 16:20:46 UTC 2013 x86_64
Build Date	Apr 4 2014 01:07:05
Server API	Apache 2.0 Handler
Virtual Directory Support	disabled
Configuration File (php.ini) Path	/etc/php5/apache2
Loaded Configuration File	/etc/php5/apache2/php.ini
Scan this dir for additional .ini files	/etc/php5/apache2/conf.d
Additional .ini files parsed	/etc/php5/apache2/conf.d/05-opcache.ini, /etc/php5/apache2/conf.d/10-pdo.ini, /etc/php5/apache2/conf.d/20-gd.ini, /etc/php5/apache2/conf.d/20-json.ini, /etc/php5/apache2/conf.d/20-mysql.ini, /etc/php5/apache2/conf.d/20-mysqli.ini, /etc/php5/apache2/conf.d/20-pdo_mysqli.ini, /etc/php5/apache2/conf.d/20-readline.ini

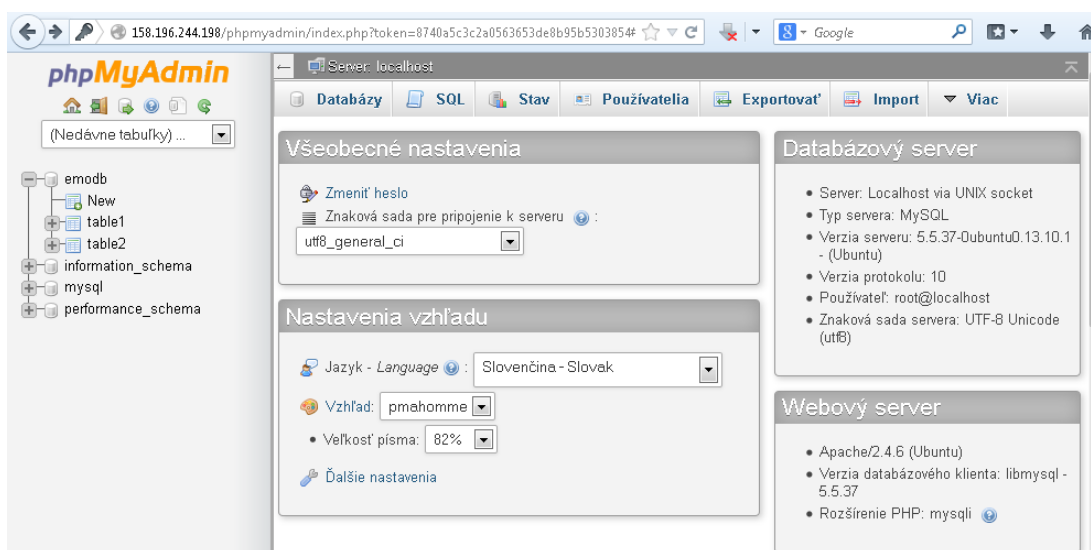
Obrázok 3.1 Výpis informácií o servery po inštalácii PHP.

3 NÁVRH A REALIZÁCIA HODNOTENIA EMÓCIÍ

Posledným krokom pred ukončením je inštalácia grafickej nadstavby k MySQL a to konkrétne phpMyAdmin. Inštalácia je opäť jednoduchá ako v predchádzajúcich krokoch, použil som tento príkaz: *apt-get install phpmyadmin*.

Následne po inštalácii som musel upraviť súbor *php.ini*, ktorý je umiestnený v zložke */etc/php5/apache2/*. Musel som nájsť riadok s výrazom *;extension=mysql.so* a odstrániť bodkočiarku z pred výrazu. A ešte bolo treba upraviť jeden súbor, pretože po zadaní adresy *http://158.196.244.198/phpmyadmin/* som dostal chybovú hlášku Error 404 File Not Found. Upravil som súbor *apache2.conf*, ktorý bol umiestnený v zložke */etc/apache2/*. Úplne na koniec súboru som vložil tento riadok: *Include /etc/phpmyadmin/apache.conf*.

Následne som reštartoval webový server a phpMyAdmin fungoval, rozhranie je zobrazené na obrázku 3.2 nižšie.



Obrázok 3.2 Rozhranie phpMyAdmin.

3.3 Návrh vzhľadu a štruktúra nástroja

Vzhľad nástroja som si po grafickej stránke chcel vybrať podľa farieb univerzity, zhruba aby tomu odpovedal. Preto som si zvolil bielo zeleno sivú kombináciu. Na vrch som potreboval nejakú titulnú grafiku, tak som si zvolil logo univerzity s týmto nápisom: *emoDBova emotional speech database of VSB-TU Ostrava*. Ukážka titulnej grafiky je nižšie na obrázku 3.3.



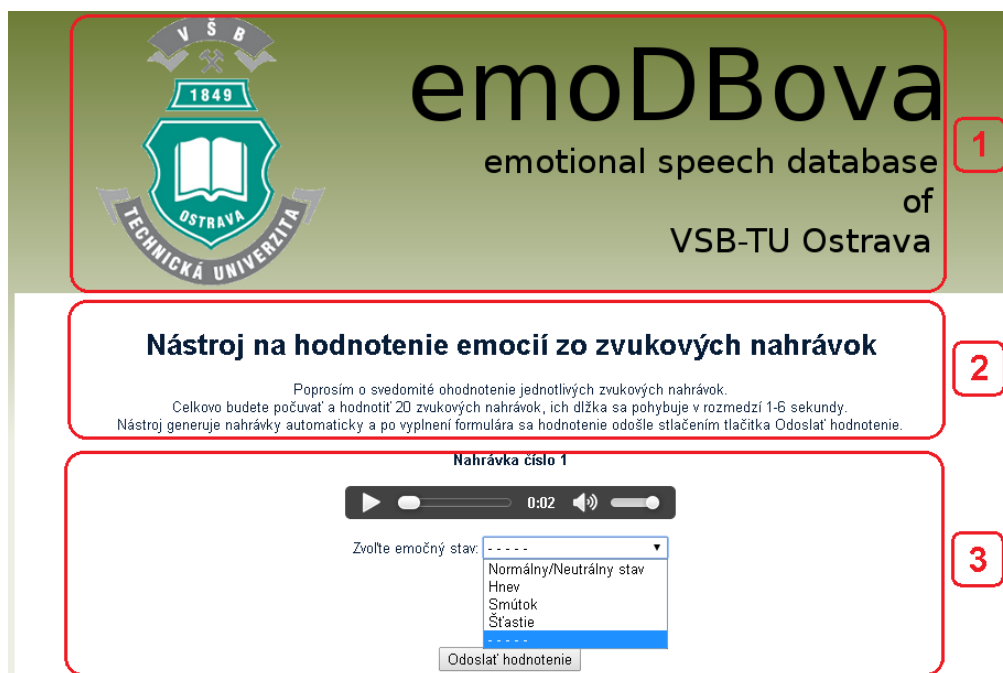
emoDBova
emotional speech database
of
VSB-TU Ostrava

Obrázok 3.3 Titulná stránka k nástroja na evaluáciu vzoriek.

3 NÁVRH A REALIZÁCIA HODNOTENIA EMÓCIÍ

Každá zo stránok nástroja je rozdelená na určité časti. Prvou časťou je titulná grafika, ktorú som už spomenul vyššie. Druhou časťou je pozadie, ktoré dotvára celkovú grafickú stránku nástroja. Tretou časťou je text, ktorý má byť buď ako oznam alebo ako inštrukciu na vykonávanie pracovnej časti nástroja. Štvrtou časťou je už spomenutá pracovná časť, ktorá obsahuje prostriedky na vykonávanie hodnotenia vzoriek. Všetky tieto časti sú zobrazené a vyznačené na obrázku 3.4 nižšie.

Nástroj je celkovo zložený zo štyroch stránok. Prvá stránka, uvítacia, má za úlohu privítať hodnotiaci subjekt a oznámiť mu, že prešiel na stránku na hodnotenie emočných vzoriek a ako má ďalej pokračovať. Druhá stránka nástroja, hodnotiacia stránka, má za úlohu vysvetliť subjektu detaily o samotnom hodnotení, resp. o jeho funkčnosti. Samozrejme má za úlohu hlavne prehrávať a hodnotiť vygenerované emočné vzorky. Na stránke pod prehrávačom je umiestnený jednoduchý formulár, ktorý obsahuje päť možností. Jedna možnosť je ako úvodná a je zvolená na začiatku po načítaní stránky. Ostatné štyri predstavujú emočné stavy: neutrálny, hnev, smútok, šťastie. Táto stránka je zobrazená na obrázku 3.4. Tretou stránkou je stránka konečná. Oznamuje subjektu, že úspešne vykonal hodnotenie všetkých 20 vzoriek. Poslednou a štvrtou v poradí je chybová stránka, ktorá sa zobrazí v prípade, že hypertextový odkaz, ktorý je generovaný počas hodnotenia, by bol upravený a údaje by boli zlé. Teda stránka obsahuje algoritmus, ktorý kontroluje obsah odkazu počas hodnotenia, ale o tom nižšie v inej časti tejto kapitoly. Ostatné stránky, okrem hodnotiacej sú zobrazené na obrázku 3.5.



Obrázok 3.4 Rozloženie jednotlivých častí nástroja, hodnotiacia stránka nástroja
1 – titulná grafika, 2 – textová časť, 3 – pracovná časť.

3.4 Návrh databázy emoDB

Počas návrhu tabuliek databázy som si dával za úlohu, aby tabuľky obsahovali všetko potrebné k analyzovaniu hodnotení a zároveň, aby obsahovali všetky potrebné informácie

3 NÁVRH A REALIZÁCIA HODNOTENIA EMÓCIÍ

o vzorkách. Keďže som nechcel použiť jednu tabuľku na všetko, pretože by bola rozsiahla a menej priehľadná, tak som vytvoril dve tabuľky. Na to, v priebehu návrhu tabuliek ma napadlo, že by som mohol použiť rovnaký názov databázy, ako používa berlínska databáza a ten znie emoDB.



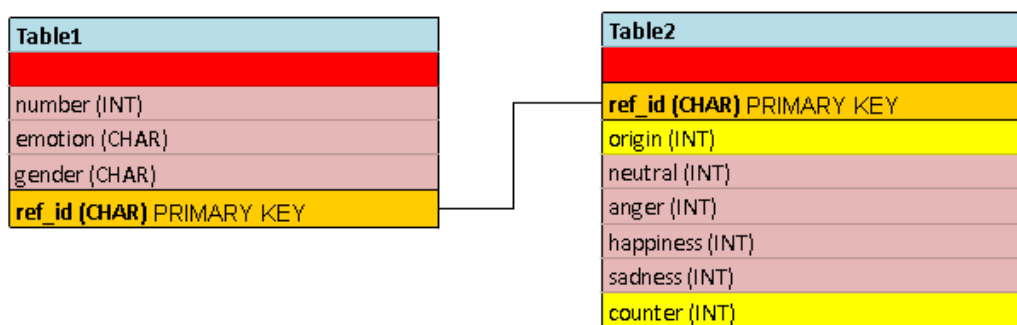
Obrázok 3.5 Všetky stránky hodnotiaceho nástroja. Hore je úvodná uvítacia stránka, uprostred je konečná stránka a dole je chybová stránka.

Tabuľka s označením *table1* je určená pre informácie o jednotlivých emočných vzorkách, obsahuje štyri stĺpce, pozri obrázok 3.6. Prvým je stĺpec s označením *number*, je to výraz použitý z anglického jazyka a znamená číslo. Tento stĺpec má predstavovať počet, koľko razy bola daná vzorka ohodnotená. Druhým stĺpec je označovaný *emotion*, čo znamená emócia po anglicky. Logicky teda sa bude do tohto stĺpca zapisovať emočný stav pre danú vzorku, ale len počiatočné písmeno z anglického výrazu. Môže to byť a pre anger, teda hnev, h pre happiness, teda radosť, n pre neutral state, teda neutrálny stav a s pre sadness, teda smútok. Tretím stĺpcom je *gender*, opäť z anglického výrazu a znamená to pohlavie. Takže tento stĺpec reprezentuje pohlavie a zapisuje sa doňho len jeden počiatočný znak z anglického výrazu pre muža alebo ženu. Buď m pre male, teda muža alebo f pre female, teda ženu. Posledný stĺpec tejto tabuľky je označovaný *ref_id* a predstavuje jedinečný názov pre označenie každej vzorky. Tento názov je popísaný vyššie v podkapitole 2.2 s nadpisom Pomenovanie nahrávok emočnej databázy. V tejto podkapitole je detailne popísaný tento jedinečný názov.

Druhou tabuľkou je tabuľka s názvom *table2*, pozri obrázok 3.6, ktorá je určená na ukladanie informácií o hodnotení jednotlivých vzoriek. Doplnkovou informáciou je pomocná premenná, ktorá je súčasťou cyklu na generovanie emočných vzoriek. Detailne to opíšem v ďalšej podkapitole. Takže tabuľka obsahuje celkovo 7 stĺpcov. Prvý v poradí je, už vyššie spomínaný, v prvej tabuľke obsiahnutý, *ref_id* stĺpec. Druhým je *origin*. Názov je použitý

3 NÁVRH A REALIZÁCIA HODNOTENIA EMÓCIÍ

z anglického výrazu, čo znamená pôvodný. Slúži na zaznamenávanie počtu hodnotení v prípade, že hodnotenie emočného stavom danej vzorky mojou osobou súhlasí s hodnotením subjektu. Ďalšie stĺpce, konkrétne tretí až šiesty, opäť slúžia na zaznamenávanie počtu hodnotení pre danú vzorku. Tento krát sa jedná o ukladanie počtu hodnotení danej vzorky pre každý emočný stav. Stĺpce majú označenie *neutral*, *anger*, *happiness*, *sadness*. Posledným stĺpcom je stĺpec s označením *counter*. Jedná sa o stĺpec, ktorý slúži na uloženie pomocnej premennej pre PHP skript.



Obrázok 3.6 Obe tabuľky databázy emoDB.

3.5 Návrh algoritmu nástroja

Aby nástroj správne fungoval a vykonal požadované úkony, bolo nutné vytvoriť niekoľko funkcií, resp. skriptov v skriptovacom jazyku PHP. Najskôr bolo nutné vytvoriť prepojenie medzi nástrojom a databázou, aby mohli byť výsledky hodnotenia zapísané. Druhým krokom, bolo vytvoriť algoritmus, ktorý by generoval jednotlivé nahrávky emočnej databázy pre prehrávač. Ešte predtým ako mohla byť nahrávka prehraná, bolo nutné vymyslieť spôsob, akým bude prehraná. Teda, či bude použitý plugin prehrávač alebo nejaký iný spôsob. A nakoniec bolo nutné vymyslieť algoritmus pre zápis dát z hodnotenia, resp. informácií o jednotlivých nahrávkach do databázy.

3.5.1 Algoritmus uvítacej stránky nástroja

Začal by som úvodnou stránkou nástroja. Na tejto stránke som použil skript pre pripojenie stránky k databáze a funkciu *database_load()*, ktorá slúži na načítanie pomocnej premennej, aby mohla byť použitá na generovanie nahrávok z databázy. Algoritmus pre pripojenie k databáze vyzerá nasledovne:

```
if(!$conn)
{
    $conn = mysql_connect($db_host, $login, $pass) or die('Connection with server
has not been established'. mysql_error());
    $selected_db = mysql_select_db($db_name, $conn) or die('Could not connect to
DB'. mysql_error());
}
```

3 NÁVRH A REALIZÁCIA HODNOTENIA EMÓCIÍ

Jednalo sa o použitie podmienky a dvoch preddefinovaných funkcií. Premenná *\$conn* je na začiatku definovaná ako *false*, vďaka tomu je skript nútený nadviazať spojenie s databázovým serverom pomocou funkcie *mysql_connect()*, ktorá má definované povinné vstupné parametre. Tieto parametre sú názov databázového servera, prihlasovacie meno a heslo. Mojom prípade to bolo *localhost*, *root*, *kukinko*. Ďalším krokom je použitie preddefinovanej funkcie *mysql_select_db()* na pripojenie ku konkrétnej databáze, mojom prípade to bola *emodb*. Funkcia má opäť povinné parametre, teda názov databázy a konkrétne MySQL pripojenie (čiže *emodb* a premenná *\$conn*). Pokiaľ nebolo zobrazená žiadna chyba, tak pripojenie bolo úspešne nadviazané. Samozrejme, ako obe funkcie sú ošetrené preddefinovanými funkciami *or die()* a *mysql_error()*, ktoré upozorňujú subjekt v prípade, že nastala chyba pri nadväzovaní pripojenia. Samozrejme pripájanie k serveru využívam aj na hodnotiacej a koncovej stránke.

Prejdem k ďalšej funkcii, ktorú som použil na tejto stránke, jedná sa funkciu *database_load()*. Služi na načítanie pomocnej premennej *\$count* z databázy, navýšenie premennej o hodnotu jedného hodnotiaceho cyklu, teda o hodnotu 5 a prepísanie pôvodnej hodnoty pomocnej premennej *\$count* uloženej v databáze o navýšenú hodnotu. Takto vyzerala táto funkcia:

```
function database_load()
{
    $help = mysql_query("SELECT `counter` FROM `table2` WHERE `ref_id`='n000f'");
    $help1 = mysql_fetch_array($help);
    $count = $help1['counter'];
    if ($count < 0 || $count > 50)
    {
        die('Invalid query: '. mysql_error());
    }
    $i = $count;
    $count += 5;
    if ($count < 50)
    {
        $query2 = 'UPDATE `table2` SET `counter`='.$count.' WHERE `ref_id`= "n000f"';
        $help = mysql_query("$query2");
    }
    elseif ($count == 50)
    {
        $count = 0;
        $query2 = 'UPDATE `table2` SET `counter`='.$count.' WHERE `ref_id`= "n000f"';
        $help = mysql_query("$query2");
    }
    return $i;
}
```

Funkcia si najprv načíta pomocou preddefinovanej funkcie *mysql_query()* hodnotu pomocnej premennej *\$count* z tabuľky *table2* zo stĺpca *counter* z riadku s *`ref_id`='n000f'*. Hodnotu tejto premennej je nutné pomocou ďalšej premennej extrahovať, pretože výstupom funkcie *mysql_query()* sú dáta typu *resource*. Takže pomocou funkcie *mysql_fetch_array()* je to možné,

3 NÁVRH A REALIZÁCIA HODNOTENIA EMÓCIÍ

pretože výstupom tejto je pole s dátami, ktoré boli získané pomocou query. Takže pomocná premenná *\$count* je načítaná. V ďalšej časti funkcie *database_load()* sú podmienky pre ošetrenie inej hodnoty premennej *\$count*. Kontrolujú, či má premenná správnu hodnotu, pokiaľ áno, tak je kontrolované či *\$count* je menšia ako 50 alebo rovnaká. Pokiaľ je menšia, tak je premenná uložená do databázy, avšak ak sa rovná 50, je hodnota nastavená na nulu a až potom uložená do databázy. Uloženie prebieha rovnako ako načítanie pomocou funkcie *mysql_query()*. Výstupom tejto funkcie je pôvodná hodnota pomocnej premennej, ktorá je používaná pri generovaní nahrávok v hodnotiacej stránke.

Pomocou POST formulára parametrom *action* sú na hodnotiacu stránku presunuté hodnoty pomocnej premennej *\$count*, ďalšej pomocnej premennej *\$next* nastavenej na hodnotu nula a *\$emo* nastavenej na hodnotu n ako neutrálna emócia.

3.5.2 Algoritmus hodnotiacej stránky nástroja

Prejdem na ďalšiu stránku nástroj, teda na hodnotiacu stránku. Na začiatku PHP skriptu stránky je načítanie hodnôt, ktoré boli pomocou POST formulára, resp. linku predané na hodnotiacu stránku. Tieto dáta sú uložené, v tom poradí ako boli predané z uvítacej stránky, do poľa s názvom *\$_GET[]*. Samozrejme hodnoty z tohto poľa sú načítané do dočasných premenných len v tom prípade, pokiaľ je toto pole naplnené dátami a pokiaľ sú splnené ďalšie podmienky. Ďalšími podmienkami je myslené, že načítavané hodnoty musia byť v určitom číselnom rozsahu. Pokiaľ nie sú splnené podmienky, tak dôjde k presmerovaniu na chybovú stránku. Stáť sa tak môže, pokiaľ budú zmenené hodnoty v linku alebo pokiaľ nebude načítaná hodnota pomocnej premennej *\$count* z databázy.

V ďalšej časti skriptu prebieha nadviazanie spojenia s databázovým serverom a načítanie emodb, ktoré som už spomínal vyššie v podkapitole 3.5.1. Teda by som prešiel k funkcií *generate_sample()*. Táto funkcia má na starosti generovanie názvu jednotlivých vzoriek, ktorý je ďalej použitý pre funkciu *audio()*. Vstupnými parametrami funkcie sú emócia a aktuálna hodnota vzorky, ktorá vychádza s premennej *\$count*. Funkcia *generate_sample()* vyzerá nasledovne:

```
function generate_sample ($emo, $i)
{
    $zero = "0";
    if($i < 10)
    {
        $file = $emo.$zero.$zero.$i;
    }
    else
    {
        $file = $emo.$zero.$i;
    }
    $gender = "f";
    $file1 = "$file$gender.wav";
    if(file_exists("sample/$file1"))
    {
        audio($file1);
        table1_insert($file, $emo, $gender);
    }
}
```

3 NÁVRH A REALIZÁCIA HODNOTENIA EMÓCIÍ

```
    table2_insert($file, $gender);
}
else
{
    $gender = "m";
    $file1 = "$file$gender.wav";
    audio($file1);
    table1_insert($file, $emo, $gender);
    table2_insert($file, $gender);
}
$file_name = array("$file", "$gender");
return $file_name;
}
```

Keďže má funkcia za úlohu generovať názov vzorky. Názov, ako je možné vidieť zo skriptu, má tvar *\$emo.\$zero.\$zero.\$i*, pokiaľ je splnená *\$i<10* podmienka alebo *\$emo.\$zero.\$i* pokiaľ podmienka nie je splnená. Táto časť funkcie vychádza z názvu vzorky spomenutého v podkapitole 2.2 Pomenovanie nahrávok emočnej databázy. Názov avšak stále nie je kompletný treba k nemu dodať na koniec pohlavie a samozrejme koncovku *.wav*. Toto je zabezpečené ďalšou podmienkou a preddefinovanou funkciou *file_exists()* pričom premenná *\$file1* obsahuje *\$file\$gender.wav*. Hodnota premennej *\$gender*, teda pohlavia, je nastavená na *f*, pričom pokiaľ súbor existuje, tak je spustená funkcia *audio()*. Ak súbor neexistuje, je hodnota *\$gender* nastavená na *m* a obsah premennej *\$file1* a je spustená funkcia *audio()*. V oboch prípadoch sú spustené funkcie *table1_insert()*, resp. *table2_insert()*, ktoré zabezpečujú ukladanie dát do databázy. Na konci sú premenné *\$file* a *\$gender* uložené do poľa a toto pole je výstupným parametrom tejto funkcie. Tieto údaje sú používané neskôr v skripte.

Ďalšou funkciou, ktorú spomeniem je *audio()*. Slúži na načítanie samotného prehrávača zvuku aj so samotnou nahrávkou, ktorej názov je vstupným parametrom funkcie. Pri výbere prehrávača som zvažoval dve možnosti, použiť plugin vo forme flashového prehrávača alebo využiť integrovaný HTML5 prehrávač zvuku. Použitie HTML5 mi prišlo ako jednoduchšie a istejšie riešenie, pretože načítanie prehrávača je rýchlejšie a nie je nutnosť externého načítania prehrávača ako v prípade flashu, resp. odpadá nutnosť inštalovať flashový prehrávač, ktorý môže spôsobovať problém s výkonom. Skript prehrávača je teda nasledovný :

```
function audio($file1)
{
    echo "<audio src='/sample/$file1' controls>
    <embed src='/sample/$file1' type='audio/wav' width='180' height='30'
    loop='false' autostart='false'></embed>
    </audio><br>";
}
```

Výstupom tohto kódu je jednoduchý zvukový prehrávač, ktorý je zobrazený na obrázku 3.6 nižšie. Prehrávač je tvorený len audio a embed HTML značkami. Je možné definovať viacej parametrov prehrávača, ako jeho veľkosť, automatické prehrávanie po načítaní nahrávky, resp. opakovanie prehrávania dookola, resp. konkrétny formát nahrávky a ešte viac ďalších parametrov.

3 NÁVRH A REALIZÁCIA HODNOTENIA EMÓCIÍ



Obrázok 3.7 Rozhranie HTML5 audio prehrávača.

Teraz prejdem na funkcie skriptu, ktoré majú za úlohu vkladanie dát do databázy emodb. Sú to funkcie *table1_insert()* a *table2_insert()*. Už z názvu je možné rozoznať, že prvá slúži na vkladanie dát o vzorkách do tabuľky *table1* a druhá ma na starosti vkladanie dát do *table2*. Vstupnými parametrami funkcií sú v oboch prípadoch sú čiastočný názov vzorky a pohlavie, v prípade funkcie *table1_insert()* je navyše ešte vstupným parametrom aj druh emócie. Teraz si uvedieme skript funkcie *table1_insert()* ako ukážku pre obe funkcie:

```
function table1_insert($file, $emo, $gender)
{
    $query = 'SELECT `ref_id` FROM `table1` WHERE `ref_id` = "'.$file.$gender.'";
    $check = mysql_query("$query");
    $check = mysql_fetch_row($check);
    if(!$check)
    {
        $query = 'INSERT INTO `table1`(`ref_id`, `number`, `emotion`, `gender`) VALUES
        ("'.$file.$gender.'", number+1, "'.$emo.'", "'.$gender.'")';
        mysql_query("$query") or die('Invalid query: '. mysql_error());
    }
    else
    {
        $query = 'UPDATE `table1` SET `number` = number+1 WHERE `ref_id` =
        "'.$file.$gender.'";
        mysql_query("$query") or die('Invalid query: '. mysql_error());
    }
}
```

Na začiatku funkcie sa kontroluje, či informácie o danej vzorke sú už obsiahnuté v tabuľke *table1*. Pokiaľ tabuľka neobsahuje informácie o danej vzorke, tak sú do tabuľky všetky informácie vložené, tak ako je definované v podkapitole 3.4. Pokiaľ sú dáta o vzorke obsiahnuté v tabuľke, tak je len navýšená hodnota stĺpca *number* o jedna. V prípade funkcie *table2_insert()* je to podobne, avšak sa jedná o vkladanie informácií o danej vzorke do tabuľky len ak tabuľka neobsahuje tieto informácie.

Ďalšou funkciou hodnotiacej stránky je *array_item()*. Úlohou tejto funkcie je získať hodnoty s premennej formulára GET alebo POST. Vstupnými parametrami funkcie sú argument, teda či sa jedná o premennú *\$_GET* alebo *\$_POST*. Výstupným parametrom je buď null alebo hodnota na danom mieste v poli. Využitie tej funkcie vyzerá nasledovne:

```
$sample_emo = array_item($_POST, 'sample_emo');
```

Súčasťou generovania názvu vzoriek je aj zmena hodnôt premenných *\$emo* a *\$next*. Následujúci algoritmus ukazuje, ako som nastavil podmienky, aby pre každú emóciu bolo

3 NÁVRH A REALIZÁCIA HODNOTENIA EMÓCIÍ

vygenerovaných práve päť nahrávok. Ďalej som zaistil, aby sa po vygenerovaní piatich nahrávok zmenil druh emócie a pokračoval opäť odznova. Skript bol napísaný nasledovne:

```
if($i < ($j+5))
{
    $sample = generate_sample($emo,$i);
    $i++;
}
elseif($i == ($j+5))
{
    $i = $j;
    $sample = generate_sample($emo,$i);
    $i++;
}
```

Táto časť generuje hodnoty poradového čísla emócie a následne s tým číslom spúšťa už spomínanú funkciu *generate_sample()* a generuje názov súboru, ktorý má byť načítaný a prehraný. Zaisťuje, aby bolo generované hodnoty od pôvodnej hodnoty načítanej s databázy algoritmom uloženú do premennej *\$i* a *\$j* (*\$j* má uloženú počiatočnú hodnotu pri načítaní z databázy, *\$i* sa zväčšuje postupne o jedna), pozri podkapitolu 3.5.1, až po hodnotu o päť väčšiu. Keď je dosiahnutá táto hodnota, tak premennej *\$i* je znovu priradená počiatočná hodnota.

```
if($next < 5)
{
    $emo = "n";
    $next++;
}
elseif($next < 10)
{
    $emo = "a";
    $next++;
}
elseif($next < 15)
{
    $emo = "h";
    $next++;
}
elseif($next < 20)
{
    $emo = "s";
    $next++;
}
```

Táto časť skriptu zisťuje, aby bola každých päť vzoriek iná premenná a premenná *\$next* pomáha, aby bola emócia postupne menená. Tento algoritmus je teda nastavený na generovanie celkom 20 vzoriek, čo znamená štyri emócie po päť vzoriek. Keď je dosiahnutá hodnota *\$next = 20*, tak na základe tejto premennej sa presmeruje z tejto stránky na konečnú stránku.

3 NÁVRH A REALIZÁCIA HODNOTENIA EMÓCIÍ

Posledná časť skriptu slúži na odoslanie dát z hodnotenia danej vzorky. Algorismus vyzerá nasledovne:

```
if(isset($_POST['selected_emo']) && $selected_emo == $sample_emo )
{
    $query = 'UPDATE `table2` SET `origin`= origin+1 WHERE `ref_id`='
    '$sample_name.''';
    mysql_query($query) or die('Invalid query: ' . mysql_error());
    $query = 'UPDATE `table2` SET `.`.$selected_emo.`= `.$selected_emo.`+1
    WHERE `ref_id`=' '$sample_name.''';
    mysql_query($query) or die('Invalid query: ' . mysql_error());
}
elseif(isset($_POST['selected_emo']))
{
    $query = 'UPDATE `table2` SET `.`.$selected_emo.`= `.$selected_emo.`+1 WHERE
    `ref_id`=' '$sample_name.''';
    mysql_query($query) or die('Invalid query: ' . mysql_error());
}
```

Prvá časť je písaná tak, aby boli hodnoty hlasovania zapísané do tabuľky *table2* v prípade, že pole *\$_POST* obsahuje dáta na pozícií *selected_emo* a moje hodnotenie vzorky sa zhoduje s hodnotením subjektu. Vtedy sa navýši hodnota pola *origin* a emócie zvolenej hodnotiacim subjektom. V druhom prípade, sa navýši hodnota len zvolenej emócie.

3.5.3 Algorismus ostatných stránok nástroja

Predposlednou stránkou je konečná stránka. Táto stránka obsahuje len dve podstatné časti kódy. Prvou je, opäť ako u oboch predchádzajúcich stránok, časť pre nadviazanie spojenia s databázou a pripojenie k databázy *emodb*. Druhou časťou je na konci predchádzajúcej podkapitoly spomenutý skript k zapísaniu dát hodnotenie do databázy.

V poslednej stránke, teda chybovej, nie je použitý žiadny skript. Stránka je napísaná bez PHP skripta, pretože len oznamuje, že nastala nejaká chyba a nerieši nič iné.

V poslednom rade je nutné spomenúť, že na konci každého skriptu, na všetkých stránkach okrem chybovej, je funkcia *mysql_close()*. Táto funkcia slúži na ukončenie aktuálneho spojenia stránky s databázovým serverom.

4 VYHODNOTENIE VÝSLEDKOV A VIEROHODNOSTI DATABÁZY

4 VYHODNOTENIE VÝSLEDKOV A VIEROHODNOSTI DATABÁZY

V tejto kapitole sa budem venovať vyhodnoteniu výsledkov získaných subjektívnym hodnotením vzoriek emočnej databázy, ktorú som sám vytvoril. Databázu najprv vyhodnotím v zmysle zistenia pravdepodobnosti s akou boli zvolené jednotlivé emócie pre každú vzorku. Ďalej určím hodnotu vierohodnosti s akou bola zvolená emócia pre danú vzorku a určím si úroveň vierohodnosti pre každú vzorku databázy. Keď už budem mať určenú úroveň vierohodnosti, tak na základe úrovne roztriedim jednotlivé vzorky do skupín.

4.1 Vyhodnotenie dosiahnutých výsledkov databázy

Hodnotenia jednotlivých emočných vzoriek boli získané pomocou subjektov, ktoré navštevovali hodiny predmetov Úvod do komunikačných technológií, resp. predmet Bezpečnosť v komunikáciách. Ďalej hodnotili vzorky databázy aj subjekty bývajúci na internáte Vysoké školy báňskej – Technickej univerzity. Subjektmi boli študenti vo vekovom rozsahu od 18 – 26 rokov. Celkovo bolo zaznamenaných 1865 hodnotení. Keď si to prepočítam na počet hodnotiacich subjektov, tak mi to vychádza cca 93 subjektov, ktorý každý z nich ohodnotil 20 vzoriek. Približný počet som uviedol z toho dôvodu, pretože celá databáza bola ohodnotená 7 razy, avšak niektoré vzorky boli ohodnotené viac ako 10 razy. Tento rozdiel mohol byť spôsobený tým, že niektoré subjekty nevykonali hodnotenie pre danú vzorku, avšak algoritmus posunul hodnotiaci cyklus ďalej a tým vznikla medzera. Alebo mohla nastať chyba v hodnotení a ako miesto hodnoty emócie pri hodnotení bola odoslaná počiatočná hodnota formulára, ktorá sa nezaznamenala do databázy. Aby sme prešli k vyhodnoteniu získaných dát, počet hodnotení pre každý emočný stav, resp. pravdepodobnosť voľby tohto stavu je zobrazená v tabuľke 4.1, resp. v grafe 4.1.

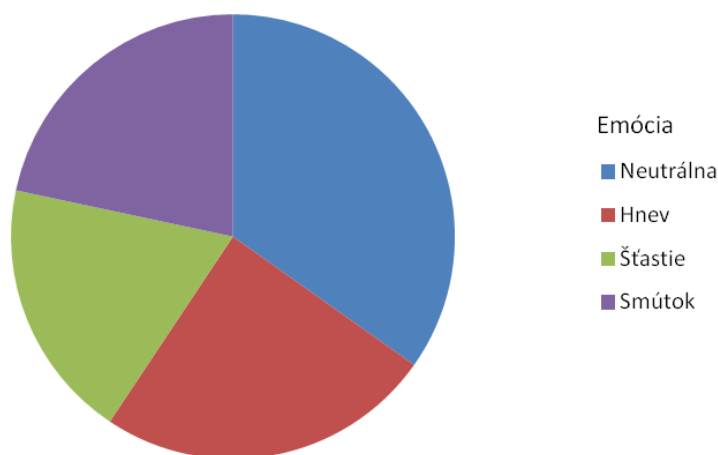
Emócia	Počet hodnotení na danú emóciu	Pravdepodobnosť s akou bola zvolená daná emócia pre 1 vzorku
Neutrálna	649	0,33
Hnev	458	0,25
Šťastie	354	0,19
Smútok	404	0,23

Tabuľka 4.1 Tabuľka obsahuje počet hodnotení a pravdepodobnosť s akou boli tieto emočné stavy zvolené.

Ako je možné vidieť z tabuľky, tak neutrálna emócia má najväčšiu prevahu v hodnotení, pretože až s 33 percentnou pravdepodobnosťou bola pri hodnotení vzoriek zvolená táto emócia, čo predstavuje jednu tretinu. Naznačuje to, že väčšia časť vzoriek má neutrálny charakter. Na grafe 4.1 je to dobre zreteľné. Nižšie pod grafom 4.1 je zobrazený krabicový graf 4.2, ktorý zobrazuje závislosť počtu hodnotení na emočný stav, pričom ako som už predtým spomínal, tak najlepšie je na tom neutrálna emócia. Ako je možno vidno z grafu, tak minimum je pri všetkých emóciách rovnaké a to nula. Maximum sa avšak líši, najväčšie je v prípade neutrálnej emócie

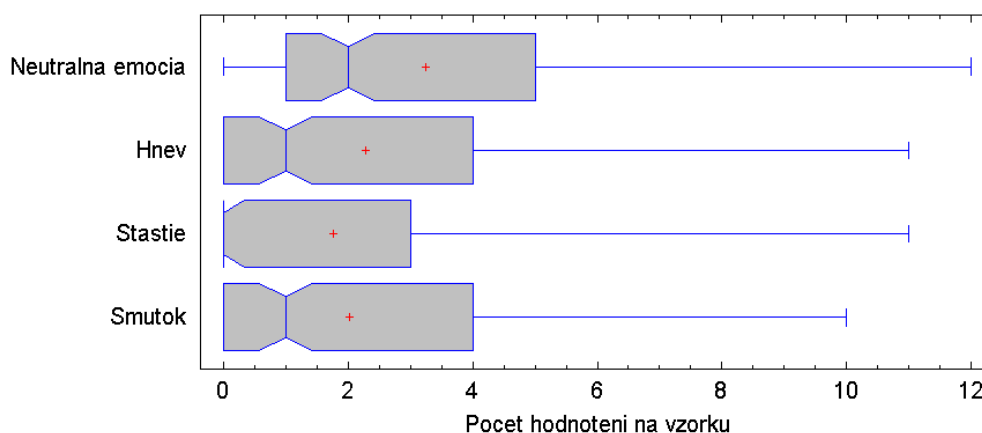
4 VYHODNOTENIE VÝSLEDKOV A VIEROHODNOSTI DATABÁZY

konkrétne 12. V prípade ostatných emócií to je 10 alebo 11. Medián je v grafe vyznačený zúženým miestom v interkvartilovom intervale. Priemer je zase vyznačený červeným znamienkom plus. Konkrétne hodnoty mediánov, priemerov a modusov vyznačených v grafe sú uvedené v tabuľke 4.2 umiestnenej nižšie na strane.



Graf 4.1 Koláčový graf znázorňuje pomer jednotlivých emočných stavov.

Posunutý kvartilový rozptyl, smerom k vyšším číslam, ako aj vyšší medián naznačuje, že hodnoty hodnotenie pre neutrálnu emóciu boli vyššie ako v prípade ostatných emócií. Z grafu je ďalej možné vyčítať, že v prípade neutrálnej emócie viac ako jedna štvrtina vzoriek mala bola hodnotených v rozsahu od 5 do 12. V prípade ostatných emócií mala štvrtina vzoriek hodnotenie od 3, resp. 4 až do 10, resp. 11. Tieto posledné kvartily mali dosť veľký rozsah v porovnaní s ostatnými, pretože v prípade hnevu, šťastia a smútku obsahovali prvé kvartily hodnoty nula, čo dokazuje aj modus v tabuľke 4.2. V prípade neutrálnej emócie je modusom číslo 2. V prípade hnevu a smútku sú hodnotenia relatívne vyrovnané, čo dokazuje aj tabuľka 4.1 resp. kvartilový rozsah, prípadne medián. Avšak podľa priemeru je zreteľné, že hnev má trochu viac hodnotení.



Graf 4.2 Krabicový graf opisujúci počet hodnotení pre jednu vzorku vzhľadom na emočný stav.

4 VYHODNOTENIE VÝSLEDKOV A VIEROHODNOSTI DATABÁZY

Emócia	Priemer	Medián	Modus
Neutrálna	3,25	2	2
Hnev	2,29	1	0
Šťastie	1,77	0	0
Smútok	2,02	1	0

Tabuľka 4.2 Hodnoty priemeru, mediánu a modusu pre všetky emočné stavy.

4.2 Určenie vierohodnosti databázy emočných vzoriek

Po ukončení hodnotenia vzoriek som tieto vzorky štatisticky vyhodnotil. Ostávalo ale ešte vykonať test vierohodnosti vzoriek. Test vierohodnosti spočíval v tom, že pre každú vzorku databázy som určil emóciu, ktorá mala najviac hodnotení. Potom som určil percento vierohodnosti tak, že som dal do pomeru najväčší počet hodnotení pre danú vzorku ku celkovému počtu hodnotení pre danú vzorku. Samozrejme som určil aj o aký emočný stav sa jednalo. V tabuľke 4.3 je ukážka toho, ako vyzerali spomenuté hodnoty pre 4 počiatkové vzorky. V tretom stĺpci je ešte aj úroveň vierohodnosti danej vzorky, ktorá je určená na základe tabuľky 4.4, ktorá obsahuje typ úrovne a hraničné hodnoty.

Vierohodnosť hodnotenia (%)	Druh vierohodnej emócie	Úroveň vierohodnosti
72,73	neutral	Low
36,36	anger	Low
45,45	neutral	Low
80,00	happiness	Medium
80,00	neutral	Medium

Tabuľka 4.3 Ukážka hodnôt vierohodnosti, druhu emócie s touto hodnotou vierohodnosti a úroveň vierohodnosti prvých piatich vzoriek. Úroveň bola stanovená na základe tabuľky 4.4.

Po stanovení hodnoty vierohodnosti som si vytvoril tri úrovne vierohodnosti, na základe ktorých rozdelím celú databázu vzoriek na tri časti. Tie úrovne sú vysoká, stredná a nízka úroveň, rozsahy úrovni sú zobrazené v tabuľke 4.4.

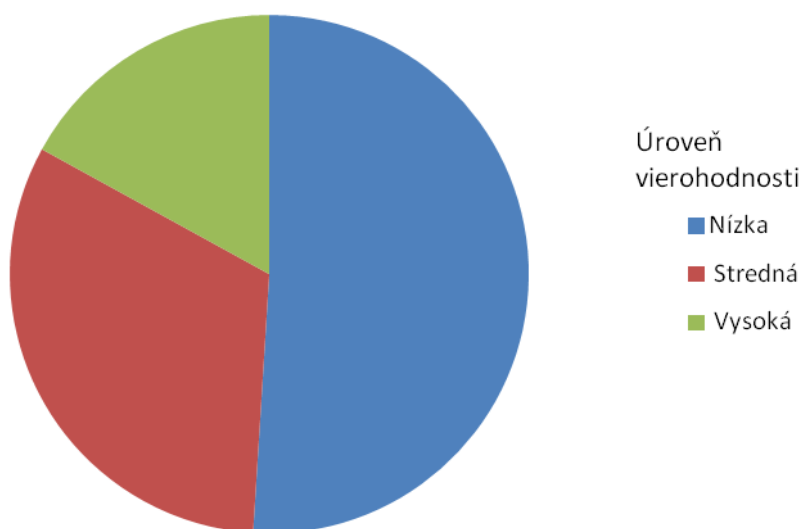
Úroveň vierohodnosti	Rozsah vierohodnosti (%)
low = nízka	< 75
medium = stredná	75 – 90
high = vysoká	> 90

Tabuľka 4.4 Tabuľka stanovuje úroveň vierohodnosti vzorky a jej rozsah.

Na základe hodnôt z tabuľky 4.4 som si vytriedil jednotlivé vzorky a vytvoril 3 skupiny vzoriek pomenovaných podľa úrovne vierohodnosti. Prvá skupina s vysokou vierohodnosťou, v rozsahu od 90 do 100 percent, obsahovala celkom 34 vzoriek s priemernou hodnotou. Táto

4 VYHODNOTENIE VÝSLEDKOV A VIEROHODNOSTI DATABÁZY

skupina vzoriek je vhodná na učenie neurálneho klasifikátora emócií doktoranda Paula Partilu. Ďalšou skupinou je skupina so strednou úrovňou vierohodnosti vzoriek. Sem patrí 64 vzoriek databázy. Táto skupina vzoriek je vhodná na tréning neurálneho klasifikátora emócií. A posledná skupina vzoriek vytvorená zo zvyšku, je tvorená zo vzoriek so zmiešanými emóciami, preto je ťažké s istotou určiť ich presný emočný stav. Na presnejšie určenie stavu, by bolo treba vykonať viac hodnotení. Na grafe 4.3 je zobrazené zastúpenie jednotlivých úrovni vierohodnosti. Všetky spomenuté úrovne a dáta súvisiace s nimi sú zobrazené v tabuľkách 4.5, 4.6 a 4.7.



Graf 4.3 Zobrazuje pomer počtu vzoriek v jednotlivých úrovniach vierohodnosti

Druh emócie	Počet vzoriek s touto emóciou	Priemerná hodnota vierohodnosti (%)
Neutrálna	4	95,99
Hnev	18	97,42
Šťastie	4	95,23
Smútok	8	97,50

Tabuľka 4.5 Skupina vzoriek s vysokou mierou vierohodnosti.

Druh emócie	Počet vzoriek s touto emóciou	Priemerná hodnota vierohodnosti (%)
Neutrálna	18	81,22
Hnev	19	82,64
Šťastie	12	80,72
Smútok	15	81,49

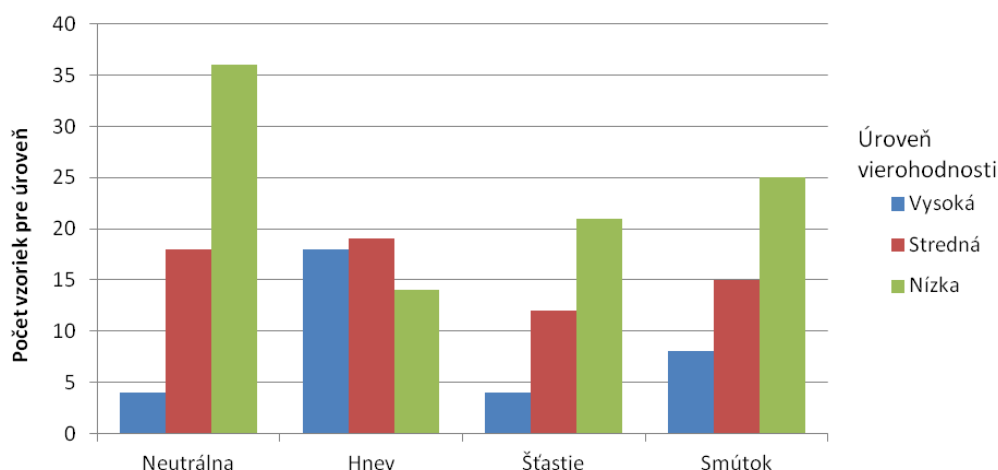
Tabuľka 4.6 Skupina vzoriek so strednou mierou vierohodnosti.

4 VYHODNOTENIE VÝSLEDKOV A VIEROHODNOSTI DATABÁZY

Druh emócie	Počet vzoriek s touto emóciou	Priemerná hodnota vierohodnosti (%)
Neutrálna	36	66,02
Hnev	14	58,21
Šťastie	21	56,07
Smútok	25	60,47

Tabuľka 4.7 Skupina vzoriek s nízkou mierou vierohodnosti.

Na grafe 4.4 je zobrazený počet vzoriek, ktoré spadajú pod jednotlivé úrovne vierohodnosti a zároveň je pre každú úroveň zobrazený aj počet vzoriek pre každý emočný stav. Z grafu je dobre možné odčítať rozdiely medzi jednotlivými úrovňami v rámci jednej emócie. V prípade neutrálnej emócie narastá počet vzoriek pre každú úroveň takmer exponenciálne. Lenže nie je to úplné tak, pretože hodnota vzoriek pre nízku úroveň je len dvojnásobná oproti strednej úrovni. V prípade šťastia a smútku môžeme konštatovať, že sa jedná lineárny medzi úrovňový nárast. Len v prípade hnevu môžeme pozorovať, že každá úroveň má približne rovnaký počet vzoriek hnevu.



Graf 4.4 Počet vzoriek pre každý emočný stav pre každú úroveň vierohodnosti.

Ked' som mal určenú úroveň vierohodnosti pre každú vzorku a roztriedil som si vzorky do úrovni, tak som si ešte v každej úrovni roztriedil vzorky podľa emócie, ktorá bola hodnotením určená. Vykonali som tak pre každý emočný stav, ktorý bol hodnotený subjektmi. Ukážka tabuľky roztriedených vzoriek pre vysokú úroveň vierohodnosti a pre emočný stav šťastie je zobrazená v tabuľke 4.8 nižšie. Obsahuje identifikačné číslo vzorky a mieru vierohodnosti pre danú vzorku.

Takže na základe určenia miery vierohodnosti, určenia úrovni vierohodnosti a roztriedenia vzoriek do jednotlivých úrovni môžem tvrdiť, že databáza vzoriek, ktorú som zostavil je, že 17 percent vzoriek je vysoko vierohodných. Môžem teda tvrdiť, že databáza je približne z 1/6 vysoko vierohodná a takmer z jednej polovice je vierohodná.

4 VYHODNOTENIE VÝSLEDKOV A VIEROHODNOSTI DATABÁZY

ID vzorky	Miera vierohodnosti (%)
109	90,91
110	100,00
135	100,00
147	90,00

Tabuľka 4.8 Ukážka roztriedených vzoriek šťastia pre vysokú úroveň vierohodnosti.

ZHODNOTENIE A ZÁVER

ZHODNOTENIE A ZÁVER

Cieľom práce bolo vytvoriť databázu vzoriek v podobe zvukových nahrávok tak, aby tieto nahrávky obsahovali reálne, teda nie hrané, emócie. Po zdĺhavej analýze a strihaní nahrávok sa mi podarilo vytvoriť databázu o 200 vzorkách o dĺžke 1-6 sekúnd na jednu vzorku. Pri vytváraní databázy sa často stávalo, že bolo počuť neželané zvuky na pozadí alebo šum, kvôli nekvalitnej nahrávke, v prípade, že sa jednalo o záznam telefonátu. Ďalším z cieľov tejto práce bolo vytvoriť nástroj na subjektívne ohodnotenie tejto databázy vzoriek za účelom zistenia emočného stavu ľudskej reči obsiahnutej vo vzorkách. Nástroj predstavuje skupinu webových stránok prepojených s MySQL databázou. Databáza slúžila nielen na ukladanie výsledkov hodnotenia, ale aj na načítanie pomocných premenných. Posledným krokom bolo vyhodnotenie výsledkov dosiahnutých hodnotení a zostavenie vierohodnej databázy vzoriek, pričom výsledná databáza obsahovala 34 vzoriek. Tieto vzorky majú vysokú vierohodnosť a sú ideálne na učenie rozpoznávania emočného stavu neurálneho klasifikátora doktoranda Partilu. Samozrejme výsledkom práce nebola len táto databáza, ale aj ďalšie dve databázy. Druhá v poradí, pokiaľ už predtým spomenutú databázu pokladáme za prvú, je databáza so strednou úrovňou vierohodnosti. Táto databáza je tvorená 64 vzorkami a je vhodná na tréning neurálneho klasifikátora. Posledná z databáz obsahuje vzorky, pre ktoré nie je možné presne určiť druh emócie. Môže sa jednať o nahrávky so zmiešanými emóciami, resp. o nahrávky, ktoré neboli dosť zreteľné.

Zámerom tohto všetkého bolo vytvoriť základ pre otvorenú databázu v tom zmysle, že by mohla byť verejne prístupná a ľudia by mohli do nej vkladať vzorky. Mohol by sa upraviť nástroj na subjektívne hodnotenie tak, že by sa pridala stránka na vkladanie vzoriek pre hodnotenie. Toto by bolo samozrejme verejne prístupné. Ďalej by tu bola možnosť pridania automatického vyhodnotenia výsledkov hodnotenia vložených vzoriek. Mohla by sa pridať aj funkcia automatického odosielania výsledkov hodnotenia.

Predstavoval by som si to tak, že subjekt by vložil do systému vzorku, ktorú by chcel ohodnotiť. Vzorka by mohla byť vložená len za predpokladu, že by spĺňala určité parametre. Subjekt by po zadaní mailovej adresy obdržal mail s informáciami a s odkazom, na ktorom si môže skontrolovať aktuálny stav hodnotenia vzorky. Subjekt by aj automaticky, v určitých časových intervaloch, dostával informácie o hodnotení. Samozrejme vzorky databázy by boli voľne dostupné na stiahnutie.

Takto nejako by vyzerala predstava ďalšieho pokračovania v tejto práci.

LITERATÚRA A ODKAZY

LITERATÚRA A ODKAZY

- [1] Západočeská univerzita v Plzni: PRAŽÁK, Aleš; RADOVÁ, Vlasta. *Úvod do praxe stínového řečníka – Proces vytváření řeči* [online]. 2014.
<<http://www.kky.zcu.cz/uploads/courses/sru/SRU2.pdf>>
- [2] PARTILA, Pavol. *Rozpoznávání emocí z řeči*. Ostrava: VŠB-TU Ostrava, 2012. 26 s.
- [3] Žilinská univerzita. *Zvuková tvorba* [online]. 2010.
<http://kt.uniza.sk/ktam/download/Studijne_materialy/Multimedia/Zvukova_tvorba/zvukova_tvorba_02_2010.pdf>
- [4] ADAMSKÝ, Aleš. *Rozpoznání emočního stavu člověka z řeči* [online]. 2009.
<https://dspace.vutbr.cz/bitstream/handle/11012/11566/BP_ALES_ADAMSKY.pdf>
- [5] NAKONEČNÝ, Milan. *Lidské emoce*. ACADEMIA, 2000. 335 s. ISBN 80-200-0763-6
- [6] ČERNÝ, Lukáš. *Rozpoznání a klasifikace emocí na základě analýzy řeči* [online]. 2009.
<https://dspace.vutbr.cz/bitstream/handle/11012/10673/DP_cerny_lukas.pdf>
- [7] Wikipedia : the free encyclopedia. *Disgust* [online]. 22. február 2003, posledná revízia 3. máj 2014. <<http://en.wikipedia.org/wiki/Disgust>>
- [8] Slovenská technická univerzita v Bratislave. *Tézy ku skúške* [online].
<http://www.ktl.elf.stuba.sk/~orgon/TEZY_skuska.pdf>
- [9] PARTILA, Pavol. *Clasification of emotion in human speech*. Ostrava: VŠB-TU Ostrava, 2013. 34 s.
- [10] VOZŇÁK, Miroslav; PARTILA, Pavol; PENHAKER, Marek; PETEREK, Tomáš; TOMALA, Karel; REZÁČ, Filip; ŠAFÁRIK, Jakub. *Emotional state and it's impact on voice authentication accuracy*. Ostrava: VŠB-TU Ostrava, 2012. 12 s.
- [11] VOZŇÁK, Miroslav; PARTILA, Pavol; MEHIC, Miralem; JAKOVLEV, Sergej. *Recognizing emotions from human speech using 2-D neural classifier and influence the selection of input parameters on its accuracy*. Ostrava: VŠB-TU Ostrava, 2013. 3 s.
- [12] VOZŇÁK, Miroslav; PARTILA, Pavol. *Speech emotions recognition using 2-D neural classifier*. Ostrava: VŠB-TU Ostrava, 2013. 10 s.

ZOZNAM PRÍLOH

ZOZNAM PRÍLOH

Príloha na DVD

- I. Diplomová práca v elektronickej podobe
- II. Materiál v podobe zvukových nahrávok, ktorý bol spracovaný na vytvorenie vzoriek
- III. Databáza vzoriek pred hodnotením
- IV. Databázy vzoriek po hodnotení subjektmi a spracovaní výsledkov
- V. Nástroj na subjektívne hodnotenie
- VI. Výsledky hodnotenia a spracované dáta